

Submitted to Econometrica

Bayesian Rapid Optimal Adaptive Design (BROAD): Method and application distinguishing models of risky choice

Debajyoti Ray¹, Daniel Golovin², Andreas Krause³ and
Colin Camerer^{4*}

October 14, 2013

BAYESIAN RAPID OPTIMAL ADAPTIVE DESIGN (BROAD): METHOD AND APPLICATION DISTINGUISHING MODELS OF RISKY CHOICE

DEBAJYOTI RAY¹, DANIEL GOLOVIN², ANDREAS KRAUSE³ AND COLIN CAMERER^{4*}

Economic surveys and experiments usually present fixed questions to respondents. Rapid computation now allows adaptively optimized questions, based on previous responses, to maximize expected information. We describe a novel method of this type introduced in computer science, and apply it experimentally to six theories of risky choice. The *EC*² method creates equivalence classes, each consisting of a true theory and its noisy-response perturbations, and chooses questions with the goal of distinguishing between equivalence classes by cutting edges connecting them. The edge-cutting information measure is adaptively submodular, which enables a provable performance bound and lazy evaluation which saves computation. The experimental data show that most subjects, making only 30 choices, can be reliably classified as choosing according to EV or two variants of prospect theory. We also consider whether subjects should and could manipulate by misreporting preferences, and find little evidence of manipulation.

KEYWORDS: JEL: C520 - Model Evaluation, Validation, and Selection; C800 - Data Collection and Data Estimation Methodology; Computer Programs: General; C910 - Design of Experiments: Laboratory, Individual;.

1. INTRODUCTION

Choices used in experimental and survey-based social science are typically developed by hunches and cumulative search for informative questions that test theories. The conventional designs that have emerged are typically a fixed set of

¹Computation and Neural Systems. California Institute of Technology. email: dray@caltech.edu

²Google, Inc., Pittsburgh, PA. email: golovin@gmail.com

³Computer Science. California Institute of Technology. email: krausea@caltech.edu

⁴Humanities and Social Sciences. California Institute of Technology. email: camerer@hss.caltech.edu, *Corresponding Author.

test questions .¹

A popular example in estimating risk preference, introduced by ? (HL, hereafter)² is a “choice list”, in which one makes a sequence of ten choices between two gambles that each have the same set of possible payoffs throughout the sequence, but in which the distribution over payoffs varies. Specifically, each gamble has two payoffs, high and low, and the probability of the high payoff is $i * 0.1$ for both gambles in the i^{th} choice in the list. Using a fixed set of items permits rapid near-replications to establish rugged stylized facts across studies.

We propose an approach, in which the sequence of choices is customized for each subject rather than fixed. The subjects themselves tell us, through their answers, the “best”(most informative) question to ask them next. We illustrate the approach using an experiment on theories of risky choice: Viz., expected utility, prospect theory, and statistical moments models (mean-variance-skewness). Our approach is abbreviated as BROAD, for **B**ayesian **R**apid **O**ptimal **A**daptive **D**esign. The BROAD method is an innovation in an old, large family of adaptive methods (described in the next subsection). The big contribution is an edge-cutting measure of information value which is adaptively submodular, which therefore provably guarantees some useful theoretical and practical properties. The method was introduced by ?, and applied here to novel economic questions, and includes new ideas about how to detect and control strategizing by subjects.

Earlier applications of adaptive methods were made in statistics (?), decision theory (?), computer-assisted testing (CAT) in psychometrics (e.g., ?) and cognitive psychology (e.g., ?), adaptive choice-based conjoint measurement in marketing (e.g., ?), “active learning” methods in computer science (e.g., ?) and machine learning (??), and management science (??).

¹Note that “test questions”, in experimental economics could include choices from budget sets ???, strategies in games, auction bids, and trading strategies. Our method could also be applied to large-scale panel surveys in which branching methods are already used (as in, “If your answer to (18) is “No” please skip to question (24)”).

²The HL paper had 1764 Google Scholar citations as of 8/27/2012.

Note that some early efforts to introduce optimal adaptive design in experimental economics (e.g., ?, ?, ?) did not gain traction. The time is now riper for BROAD methods because: Computing power is better than ever; a new method from computer science (called EC^2) applied here provides theoretical guarantees on efficient computability; and there are many new competing theories in behavioral economics which need to be efficiently compared.

In our application and all the others mentioned above, the goal is to sequentially select among a set of noisy, expensive observations (outcomes of experiments, medical tests to run to diagnose patients, etc.) in order to determine which hypothesis (theory, diagnosis, etc.) is most accurate. One way to formalize such active learning problems is *Bayesian experimental design* (?). This approach specifies prior beliefs over a set of hypotheses, as well as probabilistic assumptions about the outcomes of test choices. The goal then is to determine the correct hypothesis while minimizing the cost of the experimentation (where cost is synonymous, for present purposes, with the number of test choice questions that are asked). Unfortunately, finding an optimal sequence of test choices is not just computationally difficult (NP-hard) but is also difficult to approximate (?). Several heuristic approaches have been proposed that perform well in some specific applications, but do not have theoretical guarantees (e.g., ?); that is, there are no proofs about how costly the heuristic sequence will be compared to the optimal sequence.

In the case where observations are *noise-free*³, a simple algorithm, *generalized binary search*⁴(*GBS*), is guaranteed to be competitive with the optimal test sequence, since the expected number of tests is a factor of $O(\log n)$ (where n is the number of hypotheses) more than that of the optimal policy (?), which

³This case is known as the *Optimal Decision Tree* (ODT) problem.

⁴GBS selects test choices to maximize, in expectation over the test outcomes, the probability mass of eliminated hypotheses (i.e., those with zero posterior probability, computed with respect to the observed test outcomes). The method is myopic (or “greedy” in computer science terms) because it only searches ahead for one test at a time.

matches lower bounds up to constant factors⁵ (?).

Unfortunately, when responses are likely to be imperfect or *noisy*, as in economics experiments, the theoretical basis and performance of heuristic methods is not well understood. There are no efficient algorithms that provably achieve minimal cost, except in very restricted settings (?).⁶

In this paper, we introduce and address a general formulation of Bayesian active learning with noisy observations that we call the *Equivalence Class Determination* problem. We show that, perhaps surprisingly, generalized binary search, and some other popular algorithms, perform poorly in this setting (i.e., in simulations where the true hypothesis is known, the algorithms do not always discover the true hypothesis quickly and reliably).

However, we show that a myopic (greedy) active learning algorithm using the **Equivalence Class Edge Cutting** (EC²) objective function, can be proved to have an expected cost that is a multiple of the optimal test sequence cost. The key insight is that the EC² objective function for measuring information value satisfies *adaptive submodularity* (?). Adaptive submodularity is a natural diminishing returns property that generalizes the classical notion of submodularity to adaptive policies, and leads to a provable performance guarantee as well as a shortcut in evaluating test choices (called “lazy evaluation”).

Here are some possible *advantages* of BROAD approaches:

- The posterior distribution of all theory and parameter probabilities is quickly

⁵In other words, unless $P = NP$, no efficient algorithm can be guaranteed to find a policy whose expected number of tests is less than $C \log(n)$ times that of the optimal policy, for some absolute constant C (see ?).

⁶While there are some recent positive results in understanding the label complexity (bounds on the sample size required) of noisy active learning (??), the results depend on the assumption that the same query can be made multiple times with independent noise. This is not appropriate in our setting; if a test subject makes a mistake in evaluating a pair of gambles, presenting the exact same choice later on is likely to elicit the same mistake, which induces correlation in the noise across tests.

recomputed for each subject after each trial (because it is a necessary step in finding the optimal next question). As a result, when the experiment is over, much of the data analysis is already done.

- Because of that instant probability updating, BROAD creates a statistical parametric group portrait after each trial. The portraits show who seems most impatient, most averse to risk, most reciprocal, etc. These data could then be used to instantly form groups of different types of people to see how those differences affect group or market behavior.

- Since BROAD techniques economize on information gained per minute, they are especially useful for subject pools who have scarce time or become bored or habituated quickly. Such groups include highly-trained professionals, internet subjects who can quit, human groups such as lesion patients or children, and animals that typically make long sequences of lab choices.⁷

- The fact that the BROAD procedure generates sequences of test questions that are provably near-optimal can sharpen discourse about what different experimental designs are good and bad for. Novel BROAD designs which are unconventional should gain credibility if they have desirable informational properties. BROAD methods can also be used to judge the quality of older conventional designs⁸.

- At each trial it is easy to compute how much extra parameter precision (or model selection accuracy) is expected from asking one more question. This feature permits a cost-benefit method for optimally stopping the experiment (e.g., ?), if information benefit and cost can be compared.

In economic choice applications, there is one possible imperfection in BROAD methods: In theory, subjects might prefer to strategically manipulate their early

⁷A sensible argument is that ethical treatment of subjects requires using as little of their time as possible to learn the most ?; BROAD fits that bill.

⁸For example, retrospective analysis of influential psychology experiments on memory retention indicates that some early designs were remarkably near-optimal but others were not (?)

responses in order to get “better” (more economically valuable) future test questions. We pay special attention to the problem of manipulation, and discuss tests to detect it and methods to minimize it, within our experimental procedure. This is an important contribution of this paper, since strategic manipulation is not discussed in any detail in the most common applications (e.g. in CAT psychometric testing), and is a problem economists are well-equipped to address.

1.1. *A brief history of related adaptive methods*

The idea of optimizing an experimental design to measure parameters or test theories originated with Charles S. Peirce (?), but did not have immediate influence. Dynamic design ideas began with Wald’s influential sequential probability ratio test (?). Later contributions were made by ?, ?, and reviewed by ? and ?. D-optimal designs are optimal for estimating parameter values for a single theory. T-optimal designs optimally discriminate different theories and are more computationally challenging (cf. “landscaping”, by ?). Many of the principles originally developed under D- and T-optimality now permeate standard experimental design practice.

The idea of adaptively customizing test choices goes back 100 years to ?. It was pushed forward decades later by the US Army (?), and in psychometrics (???) in the 1960s and 70s. The most common method is computerized adaptive testing (CAT) for psychometric, intelligence (e.g., GRE, ASVAB military), health (PROMIS: ?), and psychometric measurement (see ??). CAT chooses maximally-informative items on a psychometric test, based on a subject’s previous responses, using item response theory (IRT). There are now many applications in personnel selection and professional testing⁹.

⁹E.g., ASVAB army; GRE and TOEFL; Microsoft Certification; see more examples at a clearinghouse at <http://www.psych.umn.edu/psylabs/catcentral>. Note that there are usually practical constraints imposed on large-scale CAT testing. These include balancing test content and test types and avoiding “overexposure” of highly diagnostic questions which can leak out across

Besides their statistical advantage, there is also some evidence that adaptive methods improve test-taking motivation (?), reduce boredom (?, p. 3) and reduce differences across ethnic groups (??).

Adaptive methods have also been used in neurophysiology (?), psychophysics (??), and medicine (?). In marketing, adaptive choice-based “conjoint measurement” (ACBC) is used to discover which feature configurations of products consumers like most (e.g., Sawtooth software (?). ? found modest success for ACBC in predicting actual product shares. A newer constraint satisfaction method called “poly-Q” takes a different approach, finding constrained feature weights which can explain previous choices (??). Note that improved methods will undoubtedly improve the ability to predict real behavior from responses which are hypothetical (often necessarily so). For example, ?, ?, and ? describe incentive-aligned conjoint procedures. These methods use hypothetical data from conjoint measurements or product ranks combined with random selection of a single real product for purchase (based on WTP or rank inferred from conjoint or ranks), to boost incentive-compatibility of the conjoint responses. In cognitive psychology, adaptive methods have been championed by ?, ?, ?, ???. Two applications closely related to ours are ??.

As far as we can tell from publicly available information,¹⁰ the methods above all use variants of information gain as a criterion for choosing tests (except for poly-Q). None thoroughly consider the problem of strategic manipulation.

2. MODELS

In this section we describe six models of subjective valuation of risky choices; these models will later be tested experimentally using the EC^2 method. The models are presented first in order to ground discussion of the EC^2 method in a

test-takers over time and undermine their future diagnostic value

¹⁰Details of how the Educational Testing Service uses CAT, and exactly how Sawtooth commercial ACBS software works, for example, are hard to come by.

concrete application. Note that EC^2 can be applied to many different comparison of social science theories (e.g., time preference, social preference, etc.).

The six models are widely used in different areas of social sciences and in biology. The models are: *Expected Utility with Constant Relative Risk Aversion*, *Expected Value*, *Prospect Theory*, *Cumulative Prospect Theory*, *Weighted Moments*, and *Weighted Standardized Moments*.

Choices are risky lotteries with probabilistic outcomes (a.k.a. prospects). The distribution of the payoffs in lottery L is a random variable with a range of payoffs $\mathcal{L} \equiv \{\ell_1, \dots, \ell_k\}$. Lottery L is a distribution of objectively known probabilities p_i over payoffs, $p_i \equiv \mathbb{P}[L = \ell_i]$ (with $\sum_i^k p_i = 1$). For present purposes, a payoff is an amount of money gained or lost. However, in general the lotteries can have any payoffs which are subjectively valued by numerical utilities, or by other quantities (e.g., biological fitness).

In expected utility (EU) theory, a lottery is valued by the probability-weighted sum of its possible subjective payoff utilities, $\sum_i^k p_i U(\ell_i)$ (?,?, ?).

Many different utility functions $U(\ell_i)$ have been applied and tested. We consider two popular functions $U(\ell_i)$ The first is:

Constant-Relative-Risk-Aversion (CRRA)

$$U_{CRRA}(L) = \begin{cases} \sum_i p_i (w + \ell_i)^{1-\theta} / (1-\theta) & \text{if } \theta \neq 1 \\ \sum_i p_i \log(w + \ell_i) & \text{if } \theta = 1 \end{cases}$$

Wealth w is non-negative, and in subsequent experiments we set wealth to the subject's initial endowment¹¹.

The CRRA model is used most widely for modeling risky choices (?), in economics (?, ?), psychology (?), and health (?).

The restriction of CRRA to $\theta = 0$ gives *Expected Value (EV)*

$$(2.1) \quad U_{EV}(L) = \mathbb{E}[L]$$

¹¹Wealth, w can be set somewhat arbitrarily by assuming a level of wealth for the population. In our framework, wealth can be treated as a parameter that can vary across subjects.

Under EV decision-makers are neutral toward risk— they care only about the expectation of payoff. This is an unlikely general description (especially for large payoffs) but it is a useful benchmark.

Prospect theory (PT) is a descriptive model of risky choice proposed as a more psychophysically-realistic alternative to EU (?). The key difference is that values are thought to depend on gains and losses relative to a reference point (much as sensory evaluations are generally sensitive to short-term contrast effects with previous stimuli and experiences). Reference-dependence is important because of two other properties assumed in prospect theory: First, diminishing marginal sensitivity is assumed to hold both as gains increase in magnitude (from a zero-gain point of reference) *and* as losses increase in magnitude (i.e., become more negative). This assumption implies concavity of gain utility and convexity of loss disutility (contrary to the typical EU analysis in economics, which assumes payoffs are integrated wealth positions). Second, losses are assumed to loom larger than gains, so that $v(\ell_i) < -v(-\ell_i)$ for $\ell_i > 0$.

Besides reference-dependence, PT differs from simple EU by assuming that objective probabilities are transformed into a nonlinear subjective decision weighting functions. These decision weights $\pi(p_i)$ are then used to weight reference-adjusted payoffs.

Evidence for empirical accuracy of prospect theory has come from many methods and domains¹². Nonetheless there is ongoing debate about the accuracy of prospect theory and its important features. Most prominently, there is no clear consensus on how reference points are established and change over time.¹³ As in

¹²See: Lab experiments (e.g., ?, ?); field surveys with large representative samples (?); pricing of stocks (?), and options (?, ?, ?); game shows (Post et al); and neuroscience (?, ?, ?, ?, ?, ?)

¹³A recent suggestion is that rational expectations held in the recent past about future choices and outcomes establish a reference point (?). Another important influence seems to be a sense of endowment induced, among other cues, by physical proximity (e.g., ?, ?). Note that in future designs where alternative theories of reference points make different predictions, BROAD procedures could be used to choose highly diagnostic experimental tests.

many other experiments, we assume a reference outcome at 0. That is, positive outcomes are *gains* and negative outcomes are *losses* in all choices, relative to a benchmark of a subject's wealth at the beginning of the experiment.

A CRRA power function is used for the value function v and losses are weighed by a factor λ (typically greater than one). The prospect value of a lottery is thus given by:

Prospect Theory

$$U_{PT}(L) = \sum_i v(\ell_i) \pi(p_i) \text{ where } v(\ell_i) = \begin{cases} \ell_i^\rho & \text{if } \ell_i \geq 0 \\ -\lambda(-\ell_i)^\rho & \text{if } \ell_i < 0 \end{cases}$$

Several decision weight functions $\pi(p_i)$ have been proposed (see ?, ?). In order to limit the parametric features being compared across models, we consider only the one-parameter decision weight function axiomatically derived in ?, which is $\pi(p_i) = (e^{-(\log(1/p_i))^\alpha})$. The parameter α expresses the degree of inflection: $\alpha = 1$ is linear weighting (as in EU); $\alpha < 1$ expresses the inflection hypothesized in prospect theory (?), overweighting probabilities below $p = 1/e$ and underweighting higher probabilities. The parameters $\Theta_{PT} = \{\rho, \lambda, \alpha\}$ therefore represent utility nonlinearity, loss aversion and decision weight inflection, respectively.

In the PT model just described, decision weights of all probabilities do not generally add up to one. It is well-known that this property can generate predicted violations of first-order stochastic dominance. In ? the original PT is extended by using rank-dependent weights (following ?), in a form called *Cumulative Prospect Theory (CPT)*. In CPT, a payoff is weighted by the difference between its transformed decumulative probability and the decumulative probability of the payoff ranked just below it in magnitude. Consider a gamble with two positive gain payoffs $\ell_1 > \ell_2$, with objective probabilities p_1 and p_2 and a zero payoff. Then the ranked payoffs are weighted $\pi(p_1)$, $\pi(p_2 + p_1) - \pi(p_1)$ and $1 - \pi(p_2 + p_1)$.

More generally, rank payoffs from 1 (best) to n (worst) by $\ell_1 \geq \dots \geq \ell_k \geq 0 \geq \ell_{k+1} \geq \dots \geq \ell_n$. The first k payoffs are gains and payoffs ranked $k+1$ to n are losses. Denote decision weight in the gain and loss domains by $w^+(p)$ and w^- . Then the weights are

$$\begin{aligned} \text{for } i \leq k : \pi_i^+ &= w^+(p_i + \dots + p_1) - w^+(p_{i-1} + \dots + p_1) \\ \text{for } j > k : \pi_j^- &= w^-(p_j + \dots + p_n) - w^-(p_{j+1} + \dots + p_n) \end{aligned}$$

Then the prospect value depends on ranks and the rank dependence of decision weights, as follows: *Cumulative Prospect Theory*

$$(2.2) \quad U_{CPT}(L) = \sum_{i=1}^k \pi_i^+ v(\ell_i) + \sum_{j=k+1}^n \pi_j^- v(\ell_j)$$

Besides EU variants, PT and CPT, another class of models assumes that risky choices are valued by computing and combining their statistical moments— their mean, variance, and skewness (and possibly higher moments too). Moments-based models have been most widely used in financial economics (??) and biology¹⁴.

Including skewness creates decreasing absolute risk-aversion in intermediate intervals and also allows for risk preference (?). It also appears to easily accommodate the fact that people across the world like high-skewness lottery tickets (with negative mean return; e.g. ?), and lottery-like returns appear to be valued in stocks and options (?; ?). In behavioural ecology, foraging animals have been shown to respond to mean and variance of food outcomes¹⁵.

¹⁴Criticisms of the mean-variance approach emerged early in its evolution. EU valuations can be approximated by weighting mean and variance, using a Taylor expansion argument. However, mean-variance is strictly compatible with EU-based choice only if payoff utility is quadratic or if risky choice payoffs are normally distributed (?, ?). Quadratic utility can lead to stochastic dominance valuations, and implies that absolute risk-aversion is increasing, which is unlikely (? and ?). However, analyses suggest the conclusions from mean-variance models are robust to deviations from quadratic utility and normality (??).

¹⁵? reviews the empirical and theoretical literature that investigate animals' response to vari-

Two models defined next capture the most common types of moments-based models. *Mean-Variance-Skewness (MVS)*

$$(2.3) \quad U_{MVS}(L) = w_{\mu}\mu - w_{\sigma^2}\sigma^2 + w_{\nu}\nu$$

where $\Theta_{MVS} = \{w_{\mu}, w_{\sigma^2}, w_{\nu}\}$ are the weights for the mean, variance and skewness of the lottery respectively, with $w_{\mu}, w_{\sigma^2}, w_{\nu} \geq 0$.

Since variance and skewness become much larger when payoffs are scaled up, it is also reasonable to include a standardized version of weighted moments that scales with payoff increases more gracefully: *Standardized Mean-Variance-Skewness (SMVS)*

$$(2.4) \quad U_{SMVS}(L) = w_{\mu}\mu - w_{\sigma}\sigma + w_{\tilde{\nu}}\tilde{\nu}$$

where $\Theta_{SMVS} = \{w_{\mu}, w_{\sigma}, w_{\tilde{\nu}}\}$ are the weights for the mean, standard deviation and standardized skewness of the lottery respectively, where $w_{\mu}, w_{\sigma}, w_{\tilde{\nu}} \geq 0$.

We therefore have six different models to compare. Many previous experiments tried to design choices which are most different (i.e., most diagnostic) in these different models. For example, EU and MVS approaches predict either a general aversion to payoff variance, or a preference. PT and CPT allow an aversion to variance over gains and a preference for variance over losses. It is therefore easy to design choices with gains and losses in which the theories make different predictions (see ?).

A subtler difference arises between PT/CPT, and MVS/SMVS approaches. PT and CPT predict a high subjective value for risks with low-probability high-payoff gains (due to nonlinear decision weights $\pi(p)$ on the extreme gains). MVS and SMVS also predict high subjective value for such risks because their skewness is positive. Distinguishing the two approaches therefore requires a deft ability in rate of gain. ? argue that although humans choosing monetary risks and foraging animals are averse to variance, the coefficient of variation (or relative standard deviation) is a more accurate measure of risk sensitivity.

touch, finding mixtures of high-payoff amount and high-payoff (weighted) probability such that pure skewness preference and high-payoff overweighting can be distinguished.

3. DYNAMIC OPTIMAL DESIGNS

We begin with intuition about EC^2 , then describe the method formally. It has three powerful properties which other information criteria do not have:

- Most methods are sensitive to noise.¹⁶ The empirical applications which we have in mind all feature responses with some degree of stochastic noise (i.e., people choosing according to decision rule will not make the same choice every single time from a fixed choice set). Many of the good properties of previous adaptive methods rely on deterministic, no-noise processes (so that hypotheses are immediately eliminated by a single unpredicted “mistake”). Our method EC^2 reduces the noisy case to a formally tractable noiseless case, so good performance properties carry over while noise is allowed.

- It can be proved that there is a worst-case cost bound for greedy EC^2 which depends on $\ln(1/p_{min})$ (where p_{min} is the prior probability of the least-likely theory). What does that imply? In a medical setting, for example, there could be a dangerous disease which is extremely rare (p_{min} is low) but highly contagious; and figuring out for sure whether a patient has the disease is necessary for protecting public health. This kind of diagnosis will take many tests in the worst case. However, in the empirical social science settings we are considering, the lowest p_{min} is rarely very low because it is usually adequate to consider only a modest number of possible theories and parameters (p_{min} might be on the order of 10^{-3} , for example, so that the worst-case cost multiplier is 6.9).

- Since our EC^2 information criteria is adaptively submodular, the marginal information value of a test can only go down after other tests have been done.

¹⁶Appendix B illustrates an example where information gain and value of information (VoI) perform badly.

This permits “lazy evaluation” which reduces runtime (??).

To illustrate, suppose there are 30,000 possible tests. The procedure begins by computing the value of all 30,000 tests. The highest-value test is conducted on trial 1, the result is recorded, and all model probabilities are updated. On trial 2, if a candidate new test has a computed value V , then all the tests with a value lower than V in the initial pre-trial 1 evaluation can be safely ignored (since those values may change after the trial 1 results, but can never become larger than V). This insight can be exploited, by considering tests in order of maximum previously computed values (Refer to Appendix for details).

Next we will provide some background on general methods with no noise, which sets the stage for EC². The notation is borrowed from computer science (see ?) and is designed to apply to a very wide range of domains (e.g., sensor management, medical testing, etc.). A simple visual illustration is given below (Figure 2) for the case of hypotheses about risky choice.

3.1. Bayesian Active Learning in the Noiseless Case

In the Bayesian active learning problem, we would like to distinguish among a given set of hypotheses $\mathcal{H} = \{h_1, \dots, h_n\}$ by performing tests from a set $\mathcal{T} = \{1, \dots, N\}$ of possible tests. Running test t incurs a cost of $c(t)$ and produces an outcome from a finite set of outcomes $\mathcal{X} = \{1, 2, \dots, \ell\}$.

In our setting, \mathcal{H} would be the parameter values of a models, and \mathcal{T} , the set of all possible choice pairs. \mathcal{X} is the choice observed ($\mathcal{X} = \{0, 1\}$ for a choice between two lotteries).

We let H denote the random variable which equals the true hypothesis, and model the outcome of each test t by a random variable X_t taking values in \mathcal{X} . We denote the observed outcome of test t by x_t . We further suppose we have a prior distribution P modeling our assumptions on the joint probability $P(H, X_1, \dots, X_N)$ over the hypotheses and test outcomes.

In the noiseless case, we assume that the outcome of each test is deterministic

given the true hypothesis, i.e., for each $h \in \mathcal{H}$, $P(X_1, \dots, X_N \mid H = h)$ is 0 or 1. Thus, each hypothesis h is associated with a particular vector of positive test outcomes. We assume that no two hypotheses lead to the same outcomes for all tests.¹⁷ Therefore, if we perform all tests, we can uniquely determine the true hypothesis.

However in most applications we will wish to avoid performing every possible test, as this is prohibitively expensive. Our goal is to find an adaptive policy for running tests that allows us to determine the value of H with certainty while minimizing the cost of the tests performed. Formally, a policy π (also called a conditional plan) is a partial mapping π from partial observation vectors $\mathbf{x}_{\mathcal{A}}$ to tests, specifying which test to run next (or whether we should stop testing) for any observation vector $\mathbf{x}_{\mathcal{A}}$. Hereby, $\mathbf{x}_{\mathcal{A}} \in \mathcal{X}^{\mathcal{A}}$ is a vector of outcomes indexed by a set of tests $\mathcal{A} \subseteq \mathcal{T}$ that we have performed so far¹⁸ (e.g., the set of binary choices between risky gamble pairs, or outcomes of a set of medical tests). After having made observations $\mathbf{x}_{\mathcal{A}}$, we can rule out inconsistent hypotheses. We denote the set of hypotheses consistent with an observation vector, or event Λ , by $\mathcal{V}(\Lambda) \equiv \{h \in \mathcal{H} : P(h \mid \Lambda) > 0\}$. ($\mathcal{V}(\Lambda)$ is often called the *version space* associated with Λ .)

We call a policy *feasible* if it is guaranteed to uniquely determine the correct hypothesis. That is, upon termination with observation $\mathbf{x}_{\mathcal{A}}$, it must hold that $|\mathcal{V}(\mathbf{x}_{\mathcal{A}})| = 1$. We can define the expected cost of a policy π by

$$c(\pi) \equiv \sum_h P(h) c(\mathcal{T}(\pi, h))$$

where $\mathcal{T}(\pi, h) \subseteq \mathcal{T}$ is the set of tests run by policy π in case $H = h$.

Our goal is to find a feasible policy π^* of minimum expected cost, i.e.,

$$(3.1) \quad \pi^* = \arg \min \{c(\pi) : \pi \text{ is feasible}\}$$

¹⁷That is, we implicitly assume that at least one test can be done that distinguishes any two hypotheses.

¹⁸Formally we also require that $(x_t)_{t \in \mathcal{B}} \in \text{dom}(\pi)$ and $\mathcal{A} \subseteq \mathcal{B}$, implies $(x_t)_{t \in \mathcal{A}} \in \text{dom}(\pi)$ (c.f., ?).

A policy π can be naturally represented as a decision tree T^π , and thus problem (3.1) is often called the *Optimal Decision Tree* (ODT) problem.

Unfortunately, obtaining an approximate policy π for which $c(\pi) \leq c(\pi^*) \cdot o(\log(n))$ is NP-hard (?). Hence, various heuristics are employed to solve the Optimal Decision Tree problem and its variants. Two of the most popular heuristics are to select tests by greedily maximizing the *information gain* (IG) conditioned on previous test outcomes, and *generalized binary search* (GBS). (A search is called “greedy” if it only considers the immediate effect of tests on some fixed measure of progress).

Both of the IG and GBS heuristics are greedy and after having made observations \mathbf{x}_A will select

$$t^* = \arg \max_{t \in \mathcal{T}} \Delta_{\text{Alg}}(t | \mathbf{x}_A) / c(t),$$

where $\text{Alg} \in \{\text{IG}, \text{GBS}\}$.

Here, $\Delta_{\text{IG}}(t | \mathbf{x}_A) \equiv \mathbb{H}(\mathbf{X}_T | \mathbf{x}_A) - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_A} [\mathbb{H}(\mathbf{X}_T | \mathbf{x}_A, x_t)]$ is the marginal information gain measured with respect to the Shannon entropy $\mathbb{H}(\mathbf{X}) \equiv \mathbb{E}_{\mathbf{x}}[-\log_2 P(\mathbf{x})]$, and $\Delta_{\text{GBS}}(t | \mathbf{x}_A) \equiv P(\mathcal{V}(\mathbf{x}_A)) - \sum_{x \in \mathcal{X}} P(X_t = x | \mathbf{x}_A) P(\mathcal{V}(\mathbf{x}_A, X_t = x))$ is the expected reduction in version space probability mass. Thus, both heuristics greedily chooses the test that maximizes the benefit-cost ratio, measured with respect to their particular (informational) benefit functions. The heuristics stop after running a set of tests \mathcal{A} such that $|\mathcal{V}(\mathbf{x}_A)| = 1$, i.e., once the true hypothesis has been uniquely determined. It turns out that for the (noiseless) Optimal Decision Tree problem, the IG and GBS heuristics are equivalent (?).

Interestingly, despite its greedy property GBS has been shown (?, ?, ?, ?) to obtain near-optimal expected cost: the strongest known bound is $c(\pi_{\text{GBS}}) \leq c(\pi^*) (\ln(1/p_{\min}) + 1)$ where $p_{\min} \equiv \min_{h \in \mathcal{H}} P(h)$.

Let $\mathbf{x}_S(h)$ be the unique vector $\mathbf{x}_S \in \mathcal{X}^S$ such that $P(\mathbf{x}_S | h) = 1$. The result above is proved by exploiting the fact that $f_{\text{GBS}}(S, h) \equiv 1 - P(\mathcal{V}(\mathbf{x}_S(h))) + P(h)$

is *adaptive submodular*¹⁹ and *strongly adaptively monotone*, ?. These two properties will turn out to be crucial for proving theoretical guarantees of approximation bounds for EC^2 later.

Call \mathbf{x}_A a *subvector* of \mathbf{x}_B if $A \subseteq B$ and $P(\mathbf{x}_B \mid \mathbf{x}_A) > 0$. In this case we write $\mathbf{x}_A \prec \mathbf{x}_B$. A function $f : 2^T \times \mathcal{H}$ is called *adaptive submodular* w.r.t. a distribution P , if for any $\mathbf{x}_A \prec \mathbf{x}_B$ and any test t it holds that

$$\Delta(t \mid \mathbf{x}_A) \geq \Delta(t \mid \mathbf{x}_B), \text{ where}$$

$$\Delta(t \mid \mathbf{x}_A) \equiv \mathbb{E}_H[f(\mathcal{A} \cup \{t\}, H) - f(\mathcal{A}, H) \mid \mathbf{x}_A].$$

Thus, f is *adaptive submodular* if the expected marginal benefits $\Delta(t \mid \mathbf{x}_A)$ of adding a new test t can only decrease as we gather more observations. f is called *strongly adaptively monotone* w.r.t. P if, informally, “more observations never hurt” with respect to the expected reward. Formally, for all \mathcal{A} , all $t \notin \mathcal{A}$, and all $x \in \mathcal{X}$ we require

$$\mathbb{E}_H[f(\mathcal{A}, H) \mid \mathbf{x}_A] \leq \mathbb{E}_H[f(\mathcal{A} \cup \{t\}, H) \mid \mathbf{x}_A, X_t = x].$$

The performance guarantee for IG and GBS follows from the following general result about the greedy algorithm for adaptive submodular functions (applied with $Q = 1$ and $\eta = p_{\min}$):

THEOREM 1 (Theorem 10 of ? with $\alpha = 1$) *Suppose $f : 2^T \times \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$ is adaptive submodular and strongly adaptively monotone with respect to P and there exists Q such that $f(T, h) = Q$ for all h . Let η be any value such that $f(S, h) > Q - \eta$ implies $f(S, h) = Q$ for all sets S and hypotheses h .*

Then for self-certifying²⁰ instances the adaptive greedy policy π satisfies

$$c(\pi) \leq c(\pi^*) \left(\ln \left(\frac{Q}{\eta} \right) + 1 \right).$$

¹⁹Submodularity has proved useful in many other economic contexts, e.g. ? and ?

²⁰The technical requirement that instances be *self-certifying* means that the policy will have proof that it has obtained the maximum possible objective value, Q , immediately upon doing so. It is not difficult to show that this is the case with the instances we consider in this paper. See ? for more detail.

In the following sections, we will use the concept of adaptive submodularity to provide the first approximation guarantees of test cost for Bayesian active learning with noisy observations.

3.2. The Equivalence Class Determination Problem

There are many ways to include and estimate noise in choice. The most popular method in social science is to estimate some functional specification for the noise term, and estimate the amount of noise as a parameter(s) in the general model of choice²¹. For adaptive design this approach has been pursued by Wang, Filiba and Camerer (2011). This can be problematic for two reasons:

First, with finite test vectors, it will generally be impossible to guarantee feasibility— i.e., the property that a test method will certainly locate the true hypothesis eventually— which undermines approximation guarantees. Put differently, the *version space* will not shrink rapidly because noise means that it is difficult to conclusively eliminate hypotheses.

Second, honing in on likely models both accurately and precisely requires simultaneously identifying the noise parameter and other parameters whose identification depends on noise. The results in Wang et al. (2011) are somewhat pessimistic in this regard (i.e., precision in identifying the amount of noise does not increase much over dozens of experimental trials.)

We therefore pursue a different approach. The trick is to model noise by assuming that there is random switching of “true” deterministic responses. To gain intuition, consider a simple model where tests have binary outcomes, and we know that the outcome of exactly one test, chosen uniformly at random unbeknownst to us, is flipped. If any pair of hypotheses $h \neq h'$ differs by the outcome of at least three tests, we can still uniquely determine the correct hypothesis after running all tests. In this case we can reduce the noisy active learning problem to the noiseless setting by creating N “noise-injected” copies for each hypothesis.

²¹See ?, ?, ?, ?, ?, ?, ?, ?, ?

Each noisy copy is obtained by flipping the outcome of one of the N tests. The modified prior P' would then assign mass $P'(h') = P(h)/N$ to each noisy copy h' of h . The conditional distribution $P'(\mathbf{X}_{\mathcal{T}} \mid h')$ is still deterministic (since flipping the outcome of one of the tests changes it from 0 to 1 or vice versa). Thus, each deterministic hypothesis h_i in the original problem is now associated with a set \mathcal{H}_i of deterministic hypotheses in the modified problem. However, instead of selecting tests to determine exactly which noisy copy has been realized, we only care which set of \mathcal{H}_i (all generated by flipping elements of h_i) is realized. That is, we will not necessarily be able to locate the correct noisy copy, but we can locate the family \mathcal{H}_i of noisy copies, which all spring from a single hypothesis h_i .

In the *Equivalence Class Determination problem*²², a set of hypotheses \mathcal{H} is partitioned into a set of m equivalence classes $\{\mathcal{H}_1, \dots, \mathcal{H}_m\}$ so that $\mathcal{H} = \biguplus_{i=1}^m \mathcal{H}_i$. The goal is to determine which class \mathcal{H}_i the true hypothesis lies in.

Formally, upon termination with observations $\mathbf{x}_{\mathcal{A}}$ we require that $\mathcal{V}(\mathbf{x}_{\mathcal{A}}) \subseteq \mathcal{H}_i$ for some i . As with the ODT problem, the goal is to minimize the expected cost of the tests, where the expectation is taken over the true hypothesis sampled from P . In §3.3, we will show how the ECD problem arises naturally from Bayesian experimental design problems in probabilistic models.

Given the fact that GBS performs near-optimally on the Optimal Decision Tree problem, a natural approach to solving ECD would be to run GBS until the termination condition is met. Unfortunately, and perhaps surprisingly, GBS can perform very poorly on the ECD problem. Consider an instance with a uniform prior over n hypotheses, h_1, \dots, h_n , and two equivalence classes $\mathcal{H}_1 \equiv \{h_i : 1 \leq i < n\}$ and $\mathcal{H}_2 \equiv \{h_n\}$. There are tests $\mathcal{T} = \{1, \dots, n\}$ such that $h_i(t) = 1[i = t]$, all of unit cost. Hereby, $1[\Lambda]$ is the indicator variable for event

²² ? simultaneously studied ECD, and, like us, used it to model active learning with noise (?). They developed an extension of GBS for ECD. We defer a detailed comparison of our approaches to future work.

Λ. In this case, the optimal policy only needs to select test n . However GBS may select tests $1, 2, \dots, n$ in order until running test t , where $H = h_t$ is the true hypothesis. Given our uniform prior, it takes $n/2$ tests in expectation until this happens, so that GBS costs, in expectation, $n/2$ times the optimal expected cost in this instance.

The poor performance of GBS in this instance may be attributed to its lack of consideration for the equivalence classes. Another natural heuristic would be to run the greedy IG policy, only with the entropy measured with respect to the probability distribution on *equivalence classes* rather than hypotheses. Call this policy π_{IG} . It is clearly aware of the equivalence classes, as it adaptively and myopically selects tests to reduce the uncertainty of the realized class, measured w.r.t. the Shannon entropy. However, we can show there are instances in which it costs $\Omega(n/\log(n))$ times the optimal cost, even under a uniform prior (see Appendix B for an example).

The EC^2 selection criterion.

The reason why GBS fails is because reducing the version space mass does not necessarily facilitate differentiation among the classes \mathcal{H}_i . The reason why IG fails is that there are complementarities among tests; a set of tests can be far better than the sum of its parts (so greedy test choice can fail badly because it only considers one test at a time).

Thus, we would like to optimize an objective function that encourages differentiation among classes, but lacks complementarities. We adopt an elegant idea from [?], and define weighted edges between hypotheses that we would like distinguish between. However, instead of introducing edges between arbitrary pairs of hypotheses (as done in [?]), we only introduce edges between hypotheses in different classes²³. Tests will allow us to cut edges inconsistent with their

²³We do not care to distinguish noisy copies within an equivalence class, so there is no need to create edges between these copies.

outcomes, and we aim to eliminate all inconsistent edges while minimizing the expected cost incurred. We now formalize this intuition.

Specifically, we define a set of edges $\mathcal{E} = \cup_{1 \leq i < j \leq m} \{\{h, h'\} : h \in \mathcal{H}_i, h' \in \mathcal{H}_j\}$, consisting of all (unordered) pairs of hypotheses belonging to distinct classes. These are the edges that must be *cut*, by which we mean for any edge $\{h, h'\} \in \mathcal{E}$, at least one hypothesis in $\{h, h'\}$ must be ruled out (i.e., eliminated from the version space). Hence, a test t run under true hypothesis h is said to cut edges $\mathcal{E}_t(h) \equiv \{\{h', h''\} : h'(t) \neq h(t) \text{ or } h''(t) \neq h(t)\}$. See Fig. 1 for an illustration. We define a weight function $w : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$ by $w(\{h, h'\}) \equiv P(h) \cdot P(h')$. We extend the weight function to an additive (modular) function on sets of edges in the natural manner, i.e., $w(\mathcal{E}') \equiv \sum_{e \in \mathcal{E}'} w(e)$. The objective f_{EC} that we will greedily maximize is then defined as the weight of the edges cut (EC):

$$(3.2) \quad f_{EC}(\mathcal{A}, h) \equiv w\left(\bigcup_{t \in \mathcal{A}} \mathcal{E}_t(h)\right)$$

The key insight that allows us to prove approximation guarantees for f_{EC} is that f_{EC} shares the same beneficial properties that make f_{GBS} amenable to efficient greedy optimization.

PROPOSITION 2 *The objective f_{EC} is strongly adaptively monotone and adaptively submodular.*

Based on the objective f_{EC} , we can calculate the marginal benefits for test t upon observations $\mathbf{x}_{\mathcal{A}}$ as

$$\Delta_{EC}(t | \mathbf{x}_{\mathcal{A}}) \equiv \mathbb{E}_H[f_{EC}(\mathcal{A} \cup \{t\}, H) - f_{EC}(\mathcal{A}, H) \mid \mathbf{x}_{\mathcal{A}}].$$

We call the adaptive policy π_{EC} that, after observing $\mathbf{x}_{\mathcal{A}}$, greedily selects test $t^* \in \arg \max_t \Delta_{EC}(t | \mathbf{x}_{\mathcal{A}}) / c(t)$, the EC^2 algorithm (for *equivalence class edge cutting*).

Note that these instances are self-certifying, because we obtain maximum objective value if and only if the version space lies within an equivalence class, and

the policy can certify this condition when it holds. So we can apply Theorem 1 to show EC^2 obtains a $\ln(Q/\eta) + 1$ approximation to Equivalence Class Determination. Hereby, $Q = w(\mathcal{E}) = 1 - \sum_i (P(h \in \mathcal{H}_i))^2 \leq 1$ is the total weight of all edges that need to be cut, and $\eta = \min_{e \in \mathcal{E}} w(e) \geq p_{\min}^2$ is a bound on the minimum weight among all edges. We have the following result:

THEOREM 3 *Suppose $P(h)$ is rational for all $h \in \mathcal{H}$. For the adaptive greedy policy π_{EC} implemented by EC^2 it holds that*

$$c(\pi_{\text{EC}}) \leq (2 \ln(1/p_{\min}) + 1)c(\pi^*),$$

where $p_{\min} \equiv \min_{h \in \mathcal{H}} P(h)$ is the minimum prior probability of any hypothesis, and π^* is the optimal policy for the Equivalence Class Determination problem.

3.3. Noisy Observations

We now address the case of noisy observations, using ideas from §3.2. With noisy observations, the conditional distribution $P(X_1, \dots, X_N \mid h)$ is no longer deterministic. We model the noise using an additional random variable Θ . The vector of test outcomes $\mathbf{x}_{\mathcal{T}}$ is assumed to be an arbitrary, deterministic function $\mathbf{x}_{\mathcal{T}} : \mathcal{H} \times \text{supp}(\Theta) \rightarrow \mathcal{X}^N$; hence $\mathbf{X}_{\mathcal{T}} \mid h$ is distributed as $\mathbf{x}_{\mathcal{T}}(h, \Theta_h)$ where Θ_h is distributed as $P(\theta \mid h)$. For example, there might be up to $s = |\text{supp}(\Theta)|$ ways any particular disease could manifest itself, with different patients with the same disease suffering from different symptoms.

In cases where it is always possible to identify the true hypothesis, i.e., $\mathbf{x}_{\mathcal{T}}(h, \theta) \neq \mathbf{x}_{\mathcal{T}}(h', \theta')$ for all $h \neq h'$ and all $\theta, \theta' \in \text{supp}(\Theta)$, we can reduce the problem to Equivalence Class Determination with hypotheses $\{\mathbf{x}_{\mathcal{T}}(h, \theta) : h \in \mathcal{H}, \theta \in \text{supp}(\Theta)\}$ and equivalence classes $\mathcal{H}_i \equiv \{\mathbf{x}_{\mathcal{T}}(h_i, \theta) : \theta \in \text{supp}(\Theta)\}$ for all i . Then Theorem 3 immediately yields that the approximation factor of EC^2 is at most $2 \ln(1/\min_{h, \theta} P(h, \theta)) + 1$, where the minimum is taken over all (h, θ) in the support of P . In the unit cost case, running EC^2 with a modified prior à la ?,

allows us to obtain an $\mathcal{O}(\log |\mathcal{H}| + \log |\text{supp}(\Theta)|)$ approximation factor. Note this model allows us to incorporate complex correlations in the noise terms.

However, a major challenge when dealing with noisy observations is that it is not always possible to distinguish distinct hypotheses. Even after we have run all tests, there will generally still be uncertainty about the true hypothesis, i.e., the posterior distribution $P(H \mid \mathbf{x}_T)$ obtained using Bayes' rule may still assign non-zero probability to more than one hypothesis. If so, uniquely determining the true hypothesis is not possible. Instead, we imagine that there is a set \mathcal{D} of possible *decisions* we may make after (adaptively) selecting a set of tests to perform and we must choose one (e.g., we must decide how to treat the medical patient, which scientific theory to adopt, or which classifier to use, given our observations). Thus our goal is to gather data to make effective decisions (?). Formally, for any decision $d \in \mathcal{D}$ we take, and each realized hypothesis h , we incur some loss $\ell(d, h)$. Decision theory recommends, after observing \mathbf{x}_A , to choose the decision d^* that minimizes the *risk*, i.e., the expected loss, namely $d^* \in \arg \min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_A]$.

A natural goal in Bayesian active learning is thus to adaptively pick observations, until we are guaranteed to make the same decision (and thus incur the same expected loss) that we would have made had we run *all* tests. Thus, we can reduce the noisy Bayesian active learning problem to the ECD problem by defining the equivalence classes over all test outcomes that lead to the same minimum risk decision. Hence, for each decision $d \in \mathcal{D}$, we define

$$(3.3) \quad \mathcal{H}_d \equiv \{\mathbf{x}_T : d = \arg \min_{d'} \mathbb{E}_H[\ell(d', H) \mid \mathbf{x}_T]\}.$$

If multiple decisions minimize the risk for a particular \mathbf{x}_T , we break ties arbitrarily. Identifying the best decision $d \in \mathcal{D}$ then amounts to identifying which equivalence class \mathcal{H}_d contains the realized vector of outcomes, which is an instance of ECD.

One common approach to this problem is to myopically pick tests maximizing

the decision-theoretic *value of information* (VoI) (?): $\Delta_{\text{VoI}}(t | \mathbf{x}_{\mathcal{A}}) \equiv \min_d \mathbb{E}_H[\ell(d, H) | \mathbf{x}_{\mathcal{A}}] - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_{\mathcal{A}}}[\min_d \mathbb{E}_H[\ell(d, H) | \mathbf{x}_{\mathcal{A}}, x_t]]$. The VoI of a test t is the expected reduction in the expected loss of the best decision due to the observation of x_t . However, we can show there are instances in which such a policy pays $\Omega(n/\log(n))$ times the optimal cost, even under a uniform prior on (h, θ) and with $|\text{supp}(\Theta)| = 2$ (Appendix B contains an example). In contrast, on such instances our algorithm obtains an $\mathcal{O}(\log n)$ approximation. More generally, we have the following result for EC² as an immediate consequence of Theorem 3.

THEOREM 4 *Fix hypotheses \mathcal{H} , tests \mathcal{T} with costs $c(t)$ and outcomes in \mathcal{X} , decision set \mathcal{D} , and loss function ℓ . Fix a prior $P(H, \Theta)$ and a function $\mathbf{x}_{\mathcal{T}} : \mathcal{H} \times \text{supp}(\Theta) \rightarrow \mathcal{X}^N$ which define the probabilistic noise model. Let $c(\pi)$ denote the expected cost of π incurs to find the best decision, i.e., to identify which equivalence class \mathcal{H}_d the outcome vector $\mathbf{x}_{\mathcal{T}}$ belongs to. Let π^* denote the policy minimizing $c(\cdot)$, and let π_{EC} denote the adaptive policy implemented by EC². Then it holds that*

$$c(\pi_{EC}) \leq \left(2 \ln \left(\frac{1}{p'_{\min}} \right) + 1 \right) c(\pi^*),$$

where $p'_{\min} \equiv \min_{h \in \mathcal{H}} \{P(h, \theta) : P(h, \theta) > 0\}$.

If all tests have unit cost, by using a modified prior (?) the approximation factor can be improved to $\mathcal{O}(\log |\mathcal{H}| + \log |\text{supp}(\Theta)|)$ as in the case of Theorem 3.

3.4. A simple illustration

Intuitively, EC² chooses the test which (a) distinguishes two hypotheses, connected by an edge, which make different predictions on that test, and (b) has the largest product of those two hypotheses' probability. The method is like a sports tournament in which all pairs of challengers that make different predictions are created (edges). Pairs are ranked by the product of their probabilities of being the

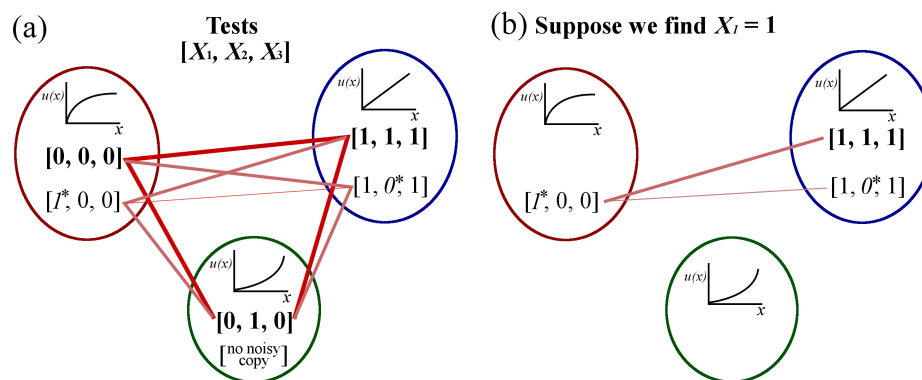


FIGURE 1.— A stylized example of Equivalence Class Edge Cutting.

“best” (i.e., most likely). The pair of contending theories with the highest probability product are then chosen to compete on a conclusive test (which will knock out one and be consistent with the other; In fact, a challenge question can maximize the sum of probability products for multiple edges, hence knocking out multiple theories at the same time). The idea is to choose the challenge which knocks out one of two theories –in our context, one of two theories that have the highest product of likelihoods of being true– as quickly as possible.

Figure 1a shows a graphic illustration of how EC^2 works. Each oval represents an equivalence class (EC) of theory predictions. In the upper left, the theory predicts choices of $[0,0,0]$ for three tests X_1, X_2, X_3 . A “noisy copy” of the theory’s prediction $[1^*,0,0]$ is randomly generated by flipping the first choice from 0 to 1. (No noisy copy is created for the bottom oval theory.) If the noise rate is 10%, then the probability of observing choices $[1^*,0,0]$ is .10 times the probability of the true predicted choices $[0,0,0]$. Edges connect all pairs of predictions between equivalence sets (although, importantly, *no edges* connect *within* an equivalence set since those predictions all stem from noisy copies of the same true theory). The thickness of the edges is a visual index of the product of the probabilities of the predictions connected by that edge. The test which can cut the sum of the thickness of multiple edges is chosen.

Suppose the test X_1 is chosen. It can distinguish the true (noiseless) predictions in left red EC and in lower green EC, which predict 0 in X_1 , from the upper right blue EC which predicts 1. If the choice is actually 1, then all theories that predicted 0 are eliminated and their edges disappear, as shown in Figure 1b.

An important caveat: Note that when implemented, the EC^2 algorithm applied to the ECD will not always choose the true hypothesis. Here's an example illustrating why: Suppose there are 10 binary tests. A hypothesis h_1 is a vector of ten 0's and 1's. Now suppose we flip those numbers so that all possible vectors are generated (i.e., there are 10 one-flip noisy copies, $10!/(10-2)!2=45$ two-flip noisy copies, etc.). If we do that for another hypothesis h_2 , then one of the noisy h_1 copies will exactly match one of the h_2 copies. We cannot draw an edge between them, because no test can distinguish them. The approach can't guarantee that we will pick the true hypothesis even if we conduct all 10 tests.

Our approach is therefore to flip some small percentage of all true hypothesis vector elements (5-20%). This approach also can not guarantee perfect accuracy. The reason is that we might have failed to randomly flip a specific test choice for a true hypothesis in the way that corresponds to an actual noisy choice a person made.

The extent to which the wrong hypotheses are selected, averaged across many simulated designs, will be seen in ground-truth simulations below (Figures 2, 3, and 4). The asymptotic probability of choosing the correct hypothesis is about 90% for the reasonable noise rates we use. That is, even if all tests are conducted, the fact that noise is injected randomly means we sometimes choose a wrong hypothesis (if the bit flipping just happened to match choices). This is an unavoidable price for a realistic noise model such as this one, which allows for human error to lead to model indistinguishability in some cases.

Method	Theoretical Properties	Reference
Information Gain	Equivalent to GBS in noise free setting.	?
Value of Information	Can only be optimized in restricted settings	?
Generalized Binary Search (GBS)	Submodular in non-adaptive setting assuming conditionally independent observations	?
Random	Asymptotic convergence when all tests applied	–

TABLE I

SUMMARY OF THE EXPERIMENTAL DESIGN METHODS THAT ARE COMPARED AGAINST EC2

4. DESIGN AND EVALUATION

The goal is to adaptively select a sequence of tests to present to a subject, in order to distinguish which of the six theories best explains the subject's responses. Each choice test t is a pair of lotteries, (L_1^t, L_2^t) . Based on the theory that represents behaviour, one of the lotteries would be preferred to the other, denoted by a binary response $x_t \in \{1, 2\}$. (Stochastic response will be introduced shortly.)

Each lottery consists of three outcomes and their corresponding probabilities, $(L_i^t) = \{(x_1, p_1), (x_2, p_2), (x_3, p_3)\}$. The middle outcome $x_2 = \$5$ is fixed in our experiments. The loss outcome x_1 varied between $\{-20, -10, -5\}$ dollars and the gain outcome x_3 varied between $\{10, 20, 50\}$ dollars. The distribution (p_1, p_2, p_3) over the payoffs was varied, where the probabilities $p_i \in \{0.01, 0.1, 0.25, 0.5, 0.75, 0.9, 0.99\}$ are chosen such that $p_1 + p_2 + p_3 = 1$.

To conduct the experiment, we used a display that showed the two lotteries along with a bar graph (see Appendix A). By considering all non-identical pairs of such lotteries, we obtained the set of 31626 possible choice test (i.e. question pairs).

4.1. Ground Truth Analyses

In this section we evaluate the quality of five different design methods, which are summarized in Table I. Both Information Gain and Generalized Binary Search are adaptive submodular in the deterministic no-noise setting (where run-

ning a test long enough is guaranteed to generate the correct response), ?. In the Random method, a test was picked uniformly at random in each trial from the set of possible tests without replacement. This should asymptotically converge to the optimal decision once all tests have been conducted. Value-of-Information (VoI) criteria is defined w.r.t a decision problem to choose which among several, potentially expensive, variables to observe in order to most effectively increase the expected utility (?). Although maximizing VoI non-myopically is optimal, in all but chain-models (e.g. Markov Chains, and slight generalizations) this non-myopic optimization is extremely intractable (?), making it infeasible for our problem.

We evaluated the speed and ability of the five methods to recover the true model based on a large sample of simulated response sequences. In each sequence, a specific true model is assumed, which generates simulated responses to the sequence of test choices generated by each method. The conclusion the method draws about which model is likely to have produced the observed responses is then compared to the true (known) model. This model-recovery exercise is sometimes called “ground truth” analysis. The responses were generated using a trembling hand function, with a known probability ϵ of making the wrong choice due to noise.

We chose discrete prior parameter values for different theories based generally on previous estimates reported in empirical literature, as follows:

- The expected value (EV) model has no parameters.
- CRRA has one risk-aversion parameter, $\Theta_{CRRA} \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$.
- A range of parameter values for prospect theory has been reported in the literature (e.g., ?, Table 1). For both PT and CPT we chose $\rho \in \{0.5, 0.62, 0.74, 0.86, 0.98, 1.1\}$, loss aversion, $\lambda \in \{1, 1.5, 2, 2.5, 3.0\}$, and Prelec-style probability weighting inflection $\alpha \in \{0.4, 0.52, 0.64, 0.76, 0.88, 1.0\}$.
- The range of parameter values for moments models MVS and SMVS were not readily available in experimental literature. To choose an appropriate range of parameter values, we created artificial choice data

based on prospect theory parameters (from the last paragraph) and found MLE best-fitting MVS and SMVS parameters.²⁴ The resulting parameter ranges were: $w_\mu \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$, $w_{\sigma^2} \in \{0.05, 0.25, 0.45, 0.65, 0.85, 1.05\} \times 10^{-2}$ and $w_\nu \in \{0, 1, 2, 3, 4, 5\} \times 10^{-4}$. For SMVS the parameter ranges chosen were: $w_\mu \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$, $w_\sigma \in \{0.05, 0.15, 0.25, 0.35, 0.45\}$ and $w_{\bar{\nu}} \in \{0.0, 0.1, 0.2, 0.3, 0.4\}$. We added a slight offset to w_{σ^2} and w_σ of 0.05 and 0.0005 respectively, to distinguish the model predictions between distinct parameter values (otherwise several values would make the identical prediction over all designs). Note that this procedure actually limits the ability of the BROAD model to find questions that sharply distinguish PT and moments models, because the candidate parameters for the moments models are those which fit PT-generated choices. However, once more information about likely moments models parameters is built up (from our data and others) the entire procedure can be repeated with a different parameter range that should be even more diagnostic than the one used here.

For the ground truth analysis, each of the six model classes started with an equal overall prior belief of $1/6$, spread uniformly across all parameter value configurations in that model class. (For example, each of the 125 specific parameter configurations for PT was assigned prior $1/6$ divided by 125.) In each ground truth run, a model and parameter configuration was chosen at random. Thirty test choices were then generated using each of the five methods. In each such run, we record whether each of the five methods is accurate after each ques-

²⁴For all possible experimental designs (roughly 32,000) in our setup, we generated the responses for all possible parameter values (125) for the prospect theory model, which generated about 4 million artificial datapoints. Next, we used a binary search to find a range of good-fitting parameter values for the moments models. We started with a wide range, and estimated the maximum likelihood fit of the generated responses. Then we did a binary split, by choosing the half of the range that contained the MLE. We repeated this procedure until dividing the parameter range led to a decrease in the likelihood of the responses, i.e. some responses are better explained with the higher parameter range.

tion. Accuracy in a run is defined as assigning the highest posterior probability to the ground-truth model (for the overall model in the same class) in that run. The procedure was repeated for 1000 runs.

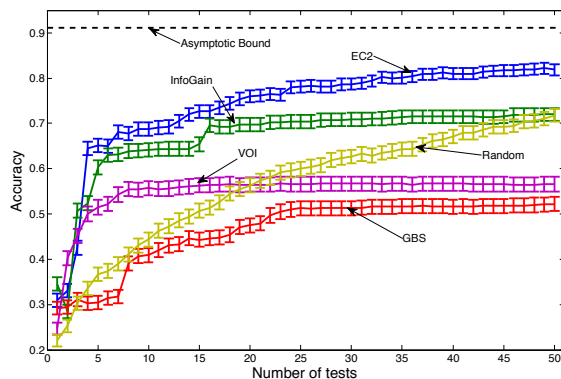


FIGURE 2.— Performance of the algorithms when probability of error, $\epsilon = 0.05$. The asymptotic bound from running all tests in random order is also shown.

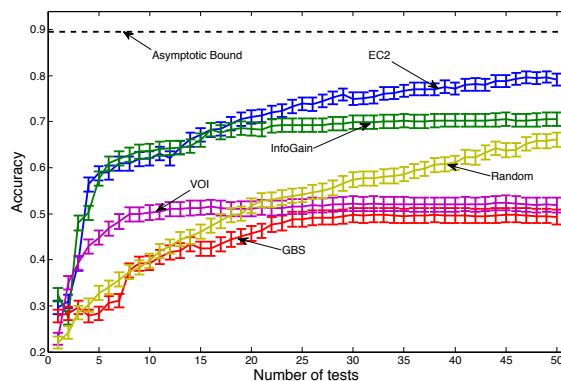


FIGURE 3.— Performance of the algorithms when probability of error, $\epsilon = 0.10$. The asymptotic bound from running all tests in random order is also shown.

Figures 2, 3, and 4 show the performance of the 5 methods for three error rates ($\epsilon = 0.05, 0.10, 0.20$).²⁵ We also plot the asymptotic performance upper

²⁵Test-retest switching rates for monetary gamble choices repeated twice are similar to these error rates, around 15% (e.g. Camerer, J Risk Uncertainty 1989).

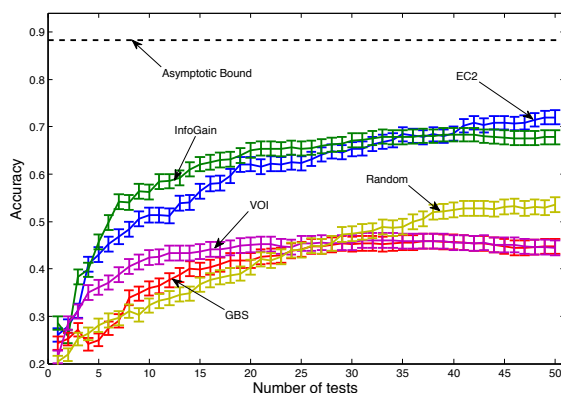


FIGURE 4.— Performance of the algorithms when probability of error, $\epsilon = 0.20$. The asymptotic bound from running all tests in random order is also shown.

bounds from running all the test choices (they are 91.09%, 89.9% and 88.2% across the three error rates).

The Figures show that VoI and GBS both perform significantly worse than Random in the presence of noise. These methods are clearly poorly-suited when responses are noisy, which most experimental social science data are likely to be. Indeed, the accuracy of the random method degrades as noise increases. EC² is noticeably more accurate than InfoGain for lower error rates, and comparable for the highest error rate $\epsilon = 0.20$. It is also evident that most of the improvement in accuracy occurs in the first 5-20 test choices; and improvement is most rapid when error is low.

These ground-truth analyses show how frequently the test selection methods get the right answer. Another performance criterion is how well the experimental setup and methods distinguish among the models. To investigate this, we plot the accuracy curves for the individual models from the ground truth process in Figure 5, for $\epsilon = 0.05$. The experimental setup we use is extremely efficient (almost 100% accuracy) at classifying SMVS, EV and CRRA. However, classification accuracy for CPT is weaker and for MVS classification is relatively poor, although significantly better than chance (which is $1/6 \approx 16.7\%$).

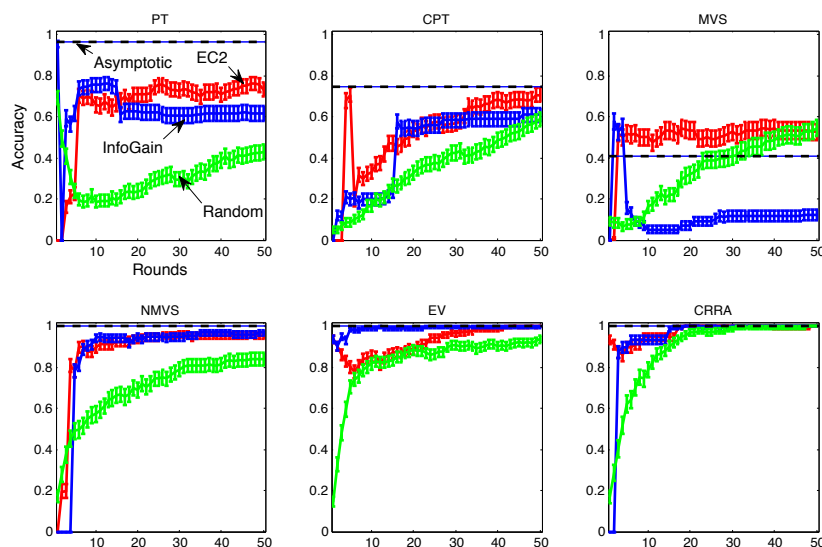


FIGURE 5.— Performance of the algorithms for individual models when probability of error, $\epsilon = 0.05$. The asymptotic bounds are also shown for recovering individual models.

We analyzed the classifications to generate a *misclassification matrix* in Table II, i.e. we looked at how the samples from the models are classified by EC². Each matrix entry shows the frequency with which a true row model is classified into each possible column model (Excellent classification would be evidenced by percentages close to 100% on the diagonal.). As pictured in Figure 5, the method correctly classifies SMVS, EV and CRRA almost perfectly. The analysis reveals where the misclassification occurs. One limit is that a significant number of PT samples (21.5%) are misclassified as CPT, and 40.6% of CPT samples are misclassified as PT. The other limit is that a large number of MVS samples (40.3%) are misclassified as EV. This is because the weight on skewness in MVS can be zero, and when the weight on variance is low enough, the choices of EV and a low-variance-weight MVS model are difficult to distinguish.

The efficiency of classification can potentially be improved through alternate experimental setups. Classification quality will also be sensitive to how the pa-

–	PT	CPT	MVS	NMVS	EV	CRRA
PT	72.7	21.5	2.9	0	2.9	0
CPT	40.6	57.3	1.1	0	0.5	0.5
MVS	0.6	0	52.7	5.8	40.3	0.6
NMVS	0	0.6	1.3	96.2	0.6	1.3
EV	0.5	0	2.6	0.5	96.4	0
CRRA	0	0	0	0	0	100.0

TABLE II

MISCLASSIFICATION MATRIX: PERCENTAGE OF SAMPLES (ROW) THAT ARE CLASSIFIED AS ANOTHER MODEL (COLUMN) USING EC2 AFTER 50 ROUNDS. THE DIAGONAL REPRESENTS CORRECT CLASSIFICATION.

parameter ranges are chosen, to how reliably subjects make choices consistent with one model or another, and other factors. Finally, keep in mind that even with an efficient setup, the method cannot necessarily achieve perfect classification using just a small number of test choices.

5. EMPIRICAL DATA AND ANALYSIS

We ran experiments with 57 subjects in 4 separate sessions at UCLA to determine which models and parameters fit their behaviour best. We used 2 different treatments using slightly different instructions (shown in Appendix A).²⁶

Thirty test choices were created using EC². To incentivize the subjects, one of their choices was selected at random for payment. Subjects were given an initial endowment of \$20 and a \$10 show-up fee²⁷. Subjects received payment based on the lottery they picked in the selected choice, and on a random outcome from that lottery²⁸. Earnings from choices ranged from \$5 to \$50, with a mean

²⁶We followed the IRB protocols of UCLA for the conduct of human subject experiments.

²⁷The earnings from the experiment were added to the endowment. Since the minimum outcome was minus \$20, the subjects would never have to pay money, as per IRB requirements.

²⁸A 10×10 matrix of values from 1 to 100 was generated by random permutation, and subjects were asked to pick a row and column number. The number at the row and column was the random probability for the subject, which was used to realize the lottery outcome. For example, if the

of \$9.21 and median of \$5.

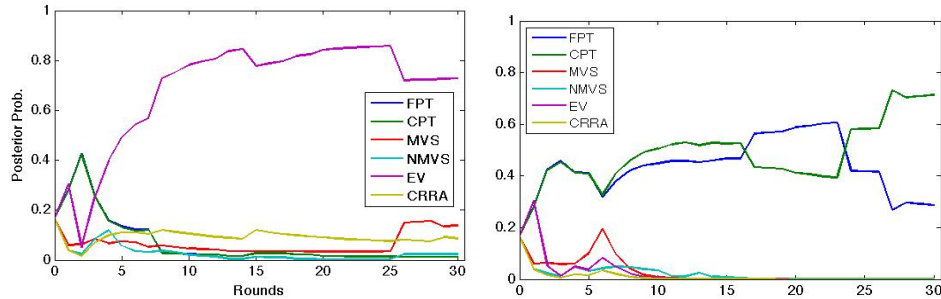


FIGURE 6.— Examples of the dynamically-estimated posterior probability over all models for two subjects. The subject on the left is rapidly classified as EV; the one on the right is less rapidly classified as CPT.

The data reveals heterogeneity in the subjects' estimated types, and the dynamics of how type estimation changes over the 30 test choices. We illustrate two examples in Figure 6. In one example the subject's type is identified very early in about 10 rounds. In the second case, the classification is evenly split between PT and CPT until the final five rounds.

Figure 7 shows stacked plots of the posterior distributions of all 57 subjects across all six models (averaged over the last five experimental trials). The stacked plot reveals high confidence in the EV classifications. The posterior mass is typically spread between the PT and CPT classifications, for those cases in which either of those models are considered the most likely. The moments models MVS and SMVS fit fewer subjects, but are typically sharply distinguished from each other. It is perhaps surprising that CRRA is never the best-fitting model.

Overall classification can be summarized by compressing the data in Figure 7. Figure 8 shows the average posterior distribution of all models after each round, aggregated over all subjects. It is clear that the posterior distribution of EV, PT and CPT are generally higher than the moments models and CRRA.

lottery picked was $L = \{(-5, 0.1), (5, 0.5), (20, 0.4)\}$, and the random number generated was 67, the subject would earn \$20.

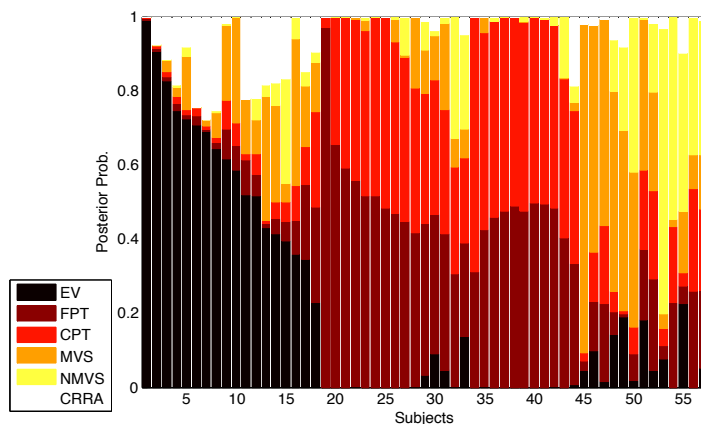


FIGURE 7.— Stacked bar plot of posterior distribution over the 6 models (averaged over the final 5 rounds) for 57 subjects. The subjects are ordered on the x-axis based on maximum posterior probability into models EV, FPT, CPT, MVS, SMVS and CRRA .

Furthermore, it appears that separation in model fit both occurs rapidly, in about five trials, and stabilizes after about 20 trials; after 20 trials the overall ranking of the model fits does not change at all.

6. IMPLEMENTATION EFFICIENCY AND CHALLENGES

Implementation speeds of BROADs are important because subjects are waiting for dynamic computation after each trial. Therefore a relevant metric of comparison is the information gained as function of time (rather than just across the number of trials). We exploit the adaptive submodular property of EC^2 to develop an *accelerated* greedy implementation that reduces computation time.

The details of the algorithm are in Appendix B. The key insight (illustrated in figure 14, Appendix B) can be gleaned from the definition of adaptive submodularity: $\Delta(t|x_B) \leq \Delta(t|x_A)$, where test $t \in \mathcal{T}$ for all $x_A \prec x_B$. The value of a test t can only decrease in future rounds as more observations are made. Initially, the values of all the tests are calculated and sorted²⁹, and the highest value test

²⁹A *Priority Queue* is often used to store sorted values for efficient calculation.

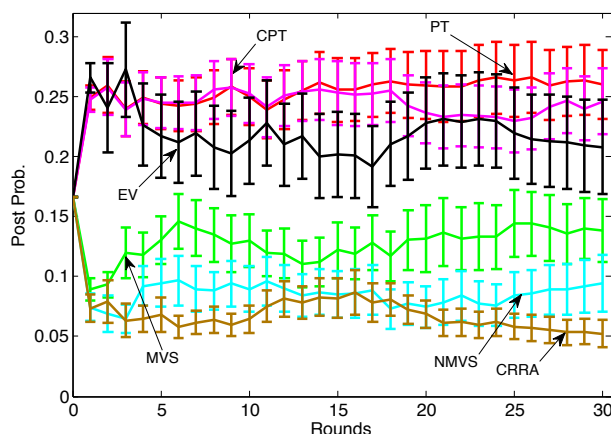


FIGURE 8.— The posterior distribution of the 6 models after each round. The aggregate mean and standard error is shown for 57 subjects.

is presented in the first round (and removed from the queue). In subsequent iterations, the value of the most valuable test is recalculated, and only the tests that had a higher value than this in the previous iteration needs to be recalculated. This saves an enormous amount of computation, and in practice could lead to orders of magnitude improvement in efficiency.

In figure 11, we plot the Information gained per unit of time for accelerated EC^2 , Information Gain and Random, assuming a response time c ($= 28$ seconds) per question. EC^2 is more efficient compared to Information Gain and Random. When the responses are immediate ($c = 0$), such as when querying a machine database, Random performs well in a short time assuming no costs for the number of queries asked. EC^2 is superior to InfoGain in both time and accuracy in this case.

7. MANIPULATION

In BROAD designs, subjects could misrepresent their true preferences in some choices *if* the following chain of conditions hold: (i) they believe that future test choices depend on previous responses; (ii) they can compute how to misrepresent

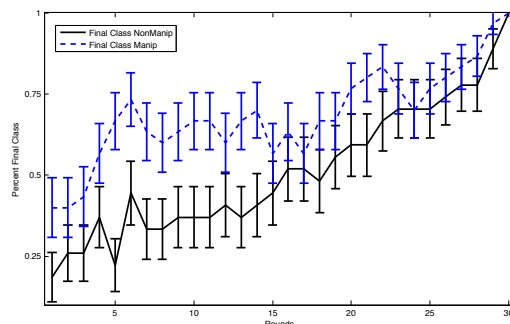


FIGURE 9.— The percentage of subjects whose current classification in a round, corresponds to final classification (in round 30). Subjects trying to manipulate the algorithm might misrepresent their true type in early rounds.

preferences in early choices to create better future choices (as evaluated by their own preferences); and (iii) the value of misrepresentation is high enough to be worthwhile.

We present arguments and evidence that misrepresentation resulting from the chain of conditions (i)-(iii) is likely to be rare. When misrepresentation does occur, it could be easily detected and is not likely to lead to wrong conclusions about revealed preferences which cannot be easily reversed.

We illustrate with a concrete example from our empirical study. For maximum fidelity, the reader should assume she has true preferences corresponding to prospect theory (PT), that was the most common type in the data.

The first test choice presented to all subjects is the same (because no choice history is used to fine-tune the test-choice selection). Subjects are asked to choose between $L_1 = \{(-10, 0.9), (5, 0), (10, 0.1)\}$ and $L_2 = \{(-20, 0.5), (5, 0.49), (50, 0.01)\}$.

What would a myopic (non-manipulating) agent choose? What would a manipulating agent choose? In post-experiment polls of this type with subjects people are not clear how manipulation would change the answer about what risky choice to pick. Even if the algorithm is known, solving the decision tree to make the optimal strategic decision is not easily computable. Here's an easier version:

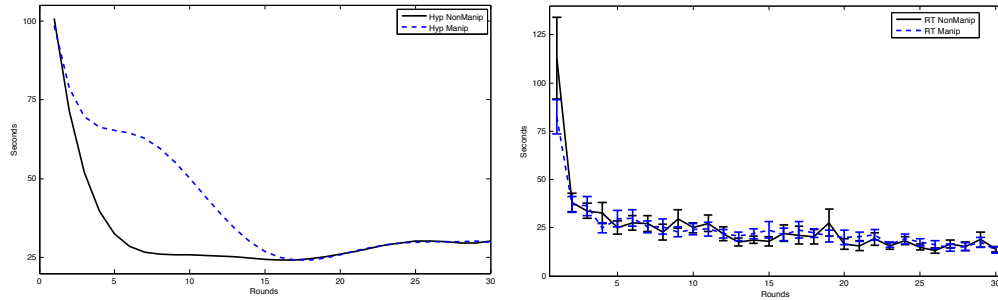


FIGURE 10.— (left) A cartoon of the hypothetical reaction times in the two treatments to test for manipulation. We expect higher reaction times when subjects are trying to manipulate. (right) The actual reaction times with standard error bars for the two treatments.

Suppose your preferences are EV. The maximum EV choice is L_2 . Should you strategically choose L_1 instead?

These examples might shake the readers' faith that manipulation is easy to figure out. But more can be said. Next we speculate about whether manipulation pays, whether it occurs, how it might be detected in experimental results, and how its effects might be muted or corrected.

◦ Does manipulation pay? To partially answer this question, it is helpful to establish an upper bound on the maximum gain from manipulation, for a particular design and player type. The upper bound on the marginal gain is likely to be low. Here's why: In later periods, it does not pay to manipulate since doing makes suboptimal immediate choices. And in early periods, manipulation is immediately costly for the same reason. So there is a natural tradeoff between the cost of manipulation in a period— the utility losses from deliberately making the wrong choices— and the future gains from improved choice sets. It could be that in a 10-period experiment, for example, manipulation is only beneficial in the first three periods. If so, the posterior probabilities computed after 10 periods might be close to the correct posteriors because they include 7 periods of non-manipulation choice data after three periods of misleading data. It is also possible

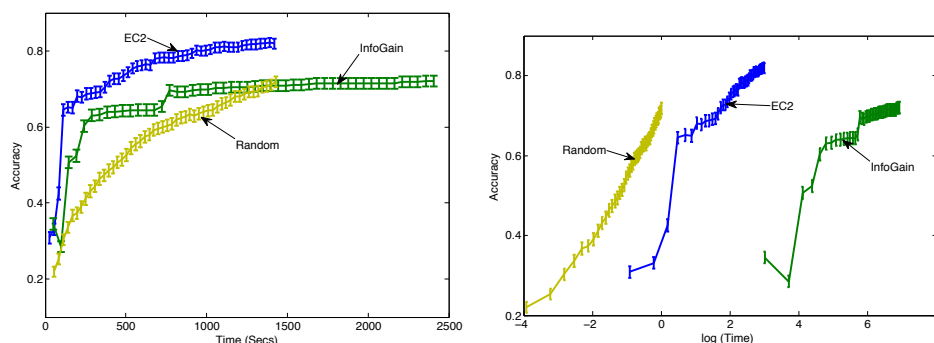


FIGURE 11.— (Left) Accuracy as a function of time assuming a response time of 28 seconds (average response time across all subjects). (Right) Accuracy as a function of logarithmic time with no response delay.

that when ranking different subjects by their risk-aversion, for example, we can recover an approximately correct ranking across subjects even if manipulation leads to biased estimates of their means.

To be more concrete, we computed the upper bound of manipulation for a *clairvoyant* agent. Specifically, consider a version of our experimental setup with 10 questions, leading to $2^{10}(= 1024)$ potential sequences of choices. We assume that the agent knows all the conditional sequences of tests, or tree, that will be generated by the EC^2 algorithm. He chooses the sequence of responses that maximize his expected earnings over all 10 trials.

We find that the earnings of a clairvoyant agent with EV preferences are only 8% higher than an agent who myopically chooses options that maximize his earning for each round. More importantly, the resulting type classification is incorrect in 68 cases, or with 6.6% error rate, which would partly be due to sampling error (Table II).

◦ Does manipulation actually happen? One way to see whether manipulation is common is to do an experiment in which one treatment group is told that an adaptive algorithm is used to generate their future choices, and it may benefit them to choose risks they do not like to induce the algorithm to offer them bet-

ter choices in the future. We created two treatments with different instructions. Subjects were told that the algorithm is adaptive and can theoretically be manipulated (Additional instructions in Appendix A). In post experiment self-reports, most subjects reported that they did not try to manipulate the system or were not able to do so. Others said they did make different choices (See online supplementary methods for self-reports). The most striking feature of the result was that the EC^2 algorithm identified more subjects in the “informed” as using expected value maximization. Some subjects said that knowing an algorithm was changing their choices led them to maximize EV. It is also notable that the subjects in the “informed” condition did not have higher average earnings (as might be expected if they are manipulating). This is explained in more detail below.

◦ Can manipulation be detected? Strategizing will typically leave clear fingerprints in the data from choices across a sequence of questions. First, in typical cases the posterior probability of the most likely model (as judged from final round result) goes up across the trials. In contrast, a manipulating respondent will appear to be one “type” in early trials, and then revert to their true “type” in later trials (as the future benefit of manipulation shrinks). This will leave a telltale pattern of posterior probabilities veering from one type to another, from earlier to later trials. We explored how diagnostic this pattern is for simulated agents who optimally misrepresent, and in actual experimental data, in figure 9. When artificial subjects are known to be manipulating, the percentage of subject classifications corresponding to the true type (as judged from final-round posteriors) is very low in early rounds. However, the opposite pattern is evident in the actual choice data.

Furthermore, if manipulation in early choices is cognitively demanding, manipulating subjects will initially have high reaction times that decrease sharply over time, as they begin to express easily-computed true preferences in later trials. However, reaction times from actual experimental data are similar in the control and cued-manipulation treatment, see Figure 10.

○ How can manipulation be limited? There are two sensible ideas about how manipulation could be limited. (Unfortunately, both of them weaken the incentives to respond truthfully to all questions that are asked.) The first remedy³⁰ is to create a space of questions Q , then sample from Q using a BROAD procedure. Subjects are told that one question will be chosen for payment *from the entire set Q* , not from the subset that were actually selected and answered. Misrepresenting preference in the first few questions in the BROAD procedure *does* lead to “better” questions being included later in the BROAD set, but it does not change at all the chance that those questions will be chosen for payment (since BROAD-selected and all other questions in Q are equally likely).

A second remedy is to make the probability with which questions are chosen for payment much higher for the earlier questions in the sequence (where the incentive for misrepresentation is highest). We will investigate numerically how much “front-loading” probability onto early questions mutes the incentive to misrepresent.

To tackle this, we ran 2 sessions, with 30 subjects at UCLA, with augmented instructions (Appendix A). Subjects were informed that an adaptive algorithm was being used, and that prior choices will affect subsequent options.

By definition, subjects would manipulate the algorithm in order to increase their earnings. The average earnings in the prior session was 9.44 dollars, with a standard deviation of 9.40, and the average earnings with the augmented instructions was 9.00 dollars with a standard deviation of 9.3. Thus the earnings from both sessions were nearly identical.

In Figure 12 we show the classifications for the two different cases to test for effects of manipulability (see Section 7 below). We cannot reject the hypothesis that the two distributions of classifications are statistically different using a 2-sided Kolmogorov-Smirnov test. We show stacked plots of the posterior distribution of the two treatments in figure 13.

³⁰This idea was suggested by Cathleen Johnson, but never published.

Furthermore, we obtained self-reports from subjects (the questions are shown in Appendix A), to find out if the participants used any strategies to game the system. The majority of the participants did not try to manipulate the algorithm. The few subjects who reported attempts to manipulate the algorithm did not come up with any effective method (summarized self-reports to be made available online).

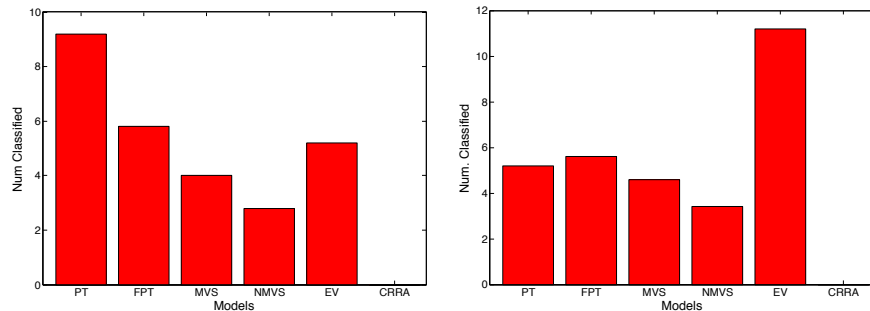


FIGURE 12.— The classifications for the two treatments to test for manipulability. Treatment 1 has 27 subjects, and treatment 2 has 30 subjects. The difference between the two distributions was not found to be statistically significant.

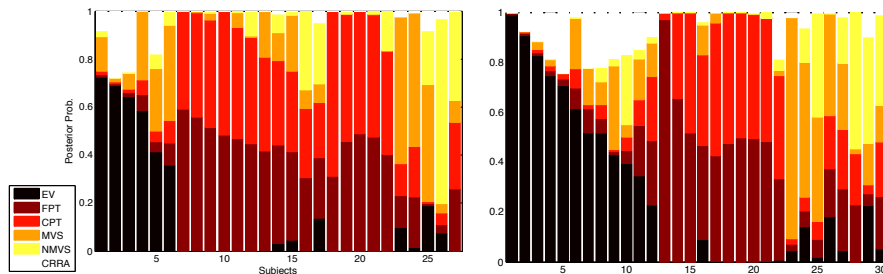


FIGURE 13.— Stacked bar plot of posterior distribution over the 6 models (averaged over the final 5 rounds) for 27 subjects in the non-manipulation condition (left) and the manipulation condition (right).

8. CONCLUSIONS AND FUTURE DIRECTIONS

This paper proposes and implements a new BROAD method for adaptively selecting which choice pairs to present to identify a distribution of economic pref-

erence parameters (and associated theories) from a known prior distribution. The novel approach here uses Equivalence Class Determination as a way to make inferences about noisy choice theories using methods developed for a classic, noiseless case. The trick is intuitively simple: A noiseless deterministic theory will make a single 0-1 choice in each choice pair. A small amount of noise is tolerated by “flipping” those 0-1 bits in a vector of choices and creating noise-injected “copies” of a target theory. Then the algorithm only tries to distinguish between an equivalence class (the target theory and all its noisy copies) and other equivalence classes.

We introduced an adaptive greedy (myopic) algorithm EC^2 , to choose questions from a design space. The algorithm measures edge strength by the product of the probabilities of two theories with similar predictions, then chooses a test which will “cut” the strongest edge (by ruling out one theory in the edge-pair and letting the other survive). The edge-cutting objective function is adaptive submodular, which leads to nice theoretical properties.

We applied EC^2 to an experimental design problem choosing among six different theories of risky choice (expectational theories, such as EU and prospect theory, and moment-weighting theories). Ground-truth simulations, which evaluate how well the algorithm approximates behavior that is known (from simulation) shows that EC^2 chooses the right theories rapidly (within 5-20 trials), and more rapidly than other theories for modest error rates.

In our empirical application, the method indicates that most subjects choosing among monetary gambles are classified (with high posterior probability) as using prospect theory evaluation, or EV.

As economists, we are also naturally concerned that if subjects understand that the algorithm is adaptive, there is a strategic incentive to misreport preferences in early trials to create better choices in latter trials. However, we argue that computing the value of misrepresentation is difficult; The value is small even when it can be computed (e.g., 8% increase in earnings), and misrepresentation

is likely to leave telltale signs in a time series of choices (high early response times and switching of posterior modes partway through the experiment) which we simply do not see in data. As a last check, in one experimental treatment we explained that the method of selecting choice questions was adaptive and that, in theory, the subjects could earn more if they exploited this property. Subjects in that treatment did not behave significantly differently and did not earn more money.

The major future direction for research is to implement these algorithms to efficiently sort through theories in many areas of empirical economics, where multiple theories have coexisted for a long time.

APPENDIX A: DETAILS OF THE EXPERIMENT

User Interface

TABLE III

The User Interface presented to the participants. The two amounts and probabilities for the 2 lotteries are shown in text on the right. The same information is displayed as a graph on the left.

Instructions

This is an experiment on the economics of decision making. If you follow the instructions and make good decisions you can earn a considerable amount of money which will be paid to you in cash. The funds have been supplied by a foundation that supports this research.

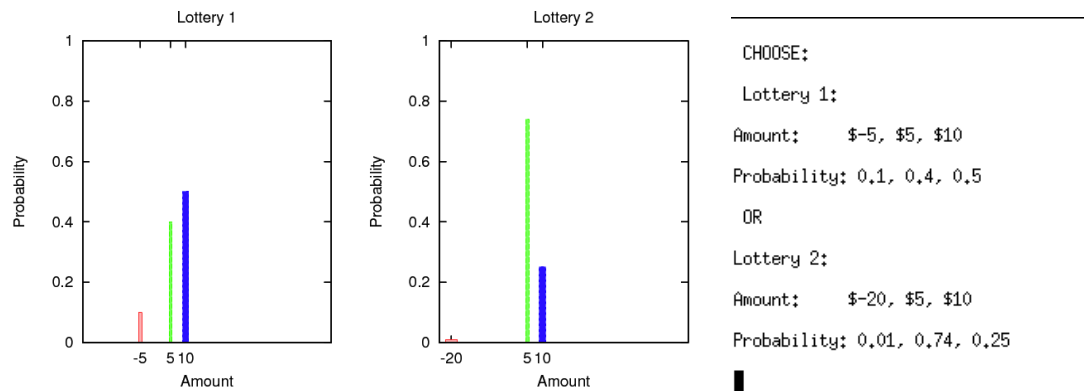
The experiment has 60 rounds. At the beginning you will begin with an endowment of \$20. In each round you will choose between one of two lotteries.

Each lottery has 3 possible outcomes:

- (i) One negative (loss) outcome of -\$5, -\$10 or -\$20. (If you lose this amount it will be subtracted from your initial \$20 endowment and you will be paid the remainder.)
- (ii) One positive (gain) outcome of \$5.
- (iii) One positive (gain) outcome of \$10, \$20 or \$50.

Each of the three possible outcomes has an associated probability of occurring. You will see a textbox with the two lottery choices described in text, on the right side of the screen. On the left side of the screen, you will see the same information displayed as a histogram. The height of each line represents the probability of that outcome occurring. The \$5 outcome is always shown in green, the higher positive outcome in blue, and the negative outcome in red.

For example, you might see the following screen:



In Lottery 1, you can lose \$5 with probability 0.1, gain \$5 with probability 0.4 or gain \$10 with probability 0.5.

In Lottery 2, you can lose \$20 with probability 0.01, gain \$5 with probability 0.74 or gain \$10 with probability 0.25.

You must close the graph display by clicking on the X (close) button. Then you can enter your choice into the textbox by pressing 1 or 2.

There is no time limit for viewing the graph and entering your choice. After you enter your choice, the screen will say REST. Another choice between two lottery options will be displayed after a couple of seconds.

After finishing all 60 rounds, we will randomly pick one of the 60 pairs of lotteries you were asked about to determine your additional earnings. Each lottery pair is equally likely to be chosen so you should treat each one as if it could be

the paid round.

To determine which of the three outcomes you receive a random integer number between 1 and 100 will determine the outcome you get. Each of those integers will be equally likely (using a procedure described next).

To use the example above, suppose you chose Lottery 1. In that lottery, the outcome -\$5 has 10% probability. Therefore, if the random number chosen is 1-10 then the outcome will be -\$5. Notice that the range 1-10 is 10% of all the numbers from 1 to 100. So the percentage of possible numbers that would give the outcome -\$5 corresponds to the probability .10.

In the lottery 1 example, the outcome will be +\$5 with probability .4, and +\$10 with probability .5. To implement these probabilities fairly, numbers from 11 to 50 will lead to the outcome +\$5 and numbers from 51-100 will lead to the outcome +\$10. Those two ranges of numbers are 40% and 50% of all the numbers 1-100, so the percentage of possible numbers corresponds to the probabilities stated in the text and shown on the graph.

Additional Instructions for Manipulability Check

These additional instructions appended to the original instructions for Manipulability Check condition:

It could help you to know that an adaptive algorithm generates the choices between two lotteries that are presented to you, in the first 30 lottery choices. That is, your previous choices determine the following questions that are presented to you, because the algorithm is trying to determine the kinds of risks you like to take and the kinds of risks you do not like to take. It is possible, mathematically, to choose lotteries you do not like in the beginning in order to lead the adaptive algorithm to create future lottery choices that you like better. If you think you can guide the algorithm in this way, by choosing lotteries you dont like to create bettery choice pairs in future trials, you are free to do so. However, your final payoff will be calculated from one

of the questions selected at random.

Quiz before the Experiment

1. There is always a possible \$5 outcome for every lottery (circle one)
TRUE FALSE
2. The red lines in the graph show the larger amounts that could be won (\$10, \$20, \$50) (circle one)
TRUE FALSE
3. If the row I choose has a random number which leads to a loss -\$20 for the lottery that I chose, I still get to keep the \$20 endowment. (circle one)
TRUE FALSE

Self-Report after Experiment

1. Did you think that the questions in the first part were generated by an adaptive algorithm (circle one)
YES NO
2. Have you figured out how the adaptive algorithm chose the subsequent questions based on your responses? If so, can you state the strategy you used to game it?
3. What strategy did you use to maximize your potential earnings in the first round? Please write a few sentences.
4. How was your strategy in the first part different from your strategy in the second part? Please write a few sentences.

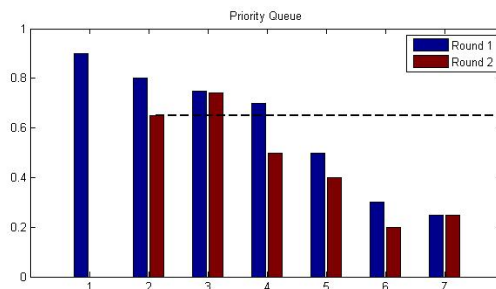


FIGURE 14.— Sorted values of tests for the accelerated EC^2 implementation.

APPENDIX B

Details of the Efficient Implementation

Figure 14 illustrates the efficient accelerated implementation of EC^2 using a Priority Queue. Initially, the tests are sorted according to their value, and the first test (number 1) is selected in the first round (and removed from the queue). In the second iterations, the value of the most valuable test (number 2) is recalculated, and only the tests that had a higher value than this in the previous iteration needs to be recalculated (number 3 and 4). The other tests (from 5) cannot have a higher value in round 2 than test number 2, and can be ignored. Algorithm 1 outlines the general procedure.

The methods were coded in Java SE 6.0. On a 2.53 GHz Intel computer with 4GB memory, the Information Gain criteria has to evaluate all 31,626 tests and takes 20.1 seconds to generate each question. The accelerated EC^2 method evaluates between 1 and up to 12 thousand tests at each iteration. By pre-computing all the values and efficient caching, the accelerated EC^2 method takes between 0.02 seconds to 6.1 seconds to generate a question. The average time is 0.4 seconds (over 1000 simulation runs of 100 rounds each).

Additional Proofs

LEMMA 5 *The objective function f of Eq. (3.2) is strongly adaptive monotone.*

Input: Set of hypotheses \mathcal{H} ; Set of tests \mathcal{T} ; prior distribution P ; function f .

begin

$\mathcal{A} \leftarrow \emptyset$; $\Psi \leftarrow \emptyset$; PriorityQueue $Q \leftarrow \text{EmptyQueue}$

foreach $t \in \mathcal{T}$ **do**

$Q.\text{insert}(t, +\infty)$

while $\exists h \neq h' : P(h \mid \mathbf{x}_{\mathcal{A}}) > 0$ and $P(h' \mid \mathbf{x}_{\mathcal{A}}) > 0$ **do**

$\delta_{max} \leftarrow -\infty$; $t_{max} \leftarrow \text{NULL}$;

while $\delta_{max} < Q.\text{maxPriority}()$ **do**

$t \leftarrow Q.\text{pop}()$;

$\delta \leftarrow \Delta(t \mid \Psi)$;

$Q.\text{insert}(t, \delta)$;

if $\delta_{max} < \delta$ **then**

$\delta_{max} \leftarrow \delta$; $t_{max} \leftarrow t$;

Set $\mathcal{A} \leftarrow \mathcal{A} \cup \{t_{max}\}$; $Q.\text{remove}(t_{max})$;

Observe outcome $x_{t_{max}}$; Set $\Psi \leftarrow \Psi \cup \{(t_{max}, x_{t_{max}})\}$

end

Algorithm 1: Algorithm for accelerated EC^2 . $Q.\text{insert}(t, \delta)$ inserts t with priority δ , $Q.\text{pop}()$ removes and returns the item with greatest priority, $Q.\text{maxPriority}()$ returns the maximum priority of the elements in Q , and $Q.\text{remove}(t)$ deletes t from Q .

PROOF: We must show that for all \mathbf{x}_A , $t \notin A$ and possible answer x for test t that

$$(8.1) \quad \mathbb{E}_H[f(A, H) \mid \mathbf{x}_A] \leq \mathbb{E}_H[f(A \cup \{t\}, H) \mid \mathbf{x}_A, X_t = x]$$

Towards this end, it is useful to notice that for all $t \in \mathcal{T}$ the function $h \mapsto \mathcal{E}_t(h)$ depends only on X_t . Hence for any \mathbf{x}_A , the function $h \mapsto f(A, h)$ is constant over realizations $\mathbf{x}_T \succ \mathbf{x}_A$, so we can define a function $g(\mathbf{x}_A)$ such that $g(\mathbf{x}_A) = \mathbb{E}_H[f(A, H) \mid \mathbf{x}_A]$ by $g(\mathbf{x}_A) := w(\bigcup_{t \in A} \mathcal{E}_t(x_t))$ where $\mathbf{x}_A = (x_t)_{t \in A}$ and $\mathcal{E}_t(x)$ is the set of edges cut by t if $X_t = x$. Note that for all $\mathbf{x}_A \prec \mathbf{x}_B$ we have $g(\mathbf{x}_A) \leq g(\mathbf{x}_B)$, since the edge weights are nonnegative. Setting $B = A \cup \{t\}$ yields Eq. (8.1) and hence implies strong adaptive monotonicity. *Q.E.D.*

LEMMA 6 *The objective function f of Eq. (3.2) is adaptive submodular for any prior with rational values.*

PROOF: We first prove the result assuming a uniform prior $P(\cdot)$, and then show how to reduce the general prior case to the uniform prior case. Hence all edges have weight $1/n^2$, where there are n hypotheses. For convenience, we also rescale our units of reward so that all edges have unit weight. (Note that f is adaptive submodular iff cf is for any constant $c > 0$.) To prove adaptive submodularity, we must show that for all $\mathbf{x}_A \prec \mathbf{x}_B$ and $t \in \mathcal{T}$, we have $\Delta(t \mid \mathbf{x}_B) \leq \Delta(t \mid \mathbf{x}_A)$. Fix t and \mathbf{x}_A , and let $\mathcal{V}(\mathbf{x}_A) := \{h : P(h \mid \mathbf{x}_A) > 0\}$ denote the version space, if \mathbf{x}_A encodes the observed outcomes. Let $n_V := |\mathcal{V}(\mathbf{x}_A)|$ be the number of hypotheses in the version space. Likewise, let $n_{i,a}(\mathbf{x}_A) := |\{h : h \in \mathcal{V}(\mathbf{x}_A, X_t = a) \cap \mathcal{H}_i\}|$, and let $n_a(\mathbf{x}_A) := \sum_{i=1}^m n_{i,a}(\mathbf{x}_A)$. Also, define $e_a(\mathbf{x}_A) := \frac{1}{2} \sum_{i \neq j} \sum_{b \neq a} n_{i,b}(\mathbf{x}_A) \cdot n_{j,b}(\mathbf{x}_A)$ to be the number of edges cut such that at t both hypotheses agree with each other but disagree with the realized hypothesis h^* , conditioning on $X_t = a$. We define a function θ of these quantities such that $\Delta(t \mid \mathbf{x}_A) = \theta(\mathbf{n}(\mathbf{x}_A), \mathbf{e}(\mathbf{x}_A))$, where $\mathbf{n}(\mathbf{x}_A)$ is the vector consisting of $n_{i,a}(\mathbf{x}_A)$ for all i and a and $\mathbf{e}(\mathbf{x}_A)$ is the vector consisting of $e_a(\mathbf{x}_A)$ for all a .

For brevity, we suppress the dependence of $\mathbf{x}_{\mathcal{A}}$ where it is unambiguous. Then, as we will explain below, θ is defined as

$$(8.2) \quad \theta(\mathbf{n}, \mathbf{e}) := \frac{1}{2} \sum_{i \neq j} \sum_{a \neq b} n_{i,a} \cdot n_{j,b} + \sum_a e_a \left(1 - \frac{n_a}{n_{\mathcal{V}}} \right)$$

Here, i and j range over all class indices, and a and b range over all possible outcomes of test t . The first term on the right-hand side counts the number of edges that will be cut by selecting test t no matter what the outcome of t is. Such edges consist of hypotheses that disagree with each other at t and, as with all edges, lie in different classes. The second term counts the expected number of edges cut by t consisting of hypotheses that agree with each other at t . Such edges will be cut by t iff they disagree with the true hypothesis h^* at t . The edges $\{h, h'\}$ with $h, h' \in \mathcal{V}(\mathbf{x}_{\mathcal{A}})$ and $P(X_t = a \mid h) = P(X_t = a \mid h') = 1$ (of which there are e_a) will be cut by t iff $X_t \neq a$. Since we assume a uniform prior, $\mathbb{P}[X_t \neq a \mid \mathbf{x}_{\mathcal{A}}] = 1 - n_a/n_{\mathcal{V}}$ for any partial realization $\mathbf{x}_{\mathcal{A}}$ with $t \notin \mathcal{A}$, hence the expected contribution of these edges to $\Delta(t \mid \mathbf{x}_{\mathcal{A}})$ is $e_a(1 - n_a/n_{\mathcal{V}})$, from whence we get the second term.

Now fix $\mathbf{x}_{\mathcal{B}} \succ \mathbf{x}_{\mathcal{A}}$. Our strategy for proving $\Delta(t \mid \mathbf{x}_{\mathcal{B}}) \leq \Delta(t \mid \mathbf{x}_{\mathcal{A}})$ is as follows. As more observations are made, the version space can only shrink, i.e. $\mathcal{V}(\mathbf{x}_{\mathcal{B}}) \subseteq \mathcal{V}(\mathbf{x}_{\mathcal{A}})$. This means that for all i and a , $n_{i,a}$ is nonincreasing, i.e., $n_{i,a}(\mathbf{x}_{\mathcal{B}}) \leq n_{i,a}(\mathbf{x}_{\mathcal{A}})$. Note we may interpret e_a as a function of the variables in $\{n_{i,a} : 1 \leq i \leq m, a \in \mathcal{X}\}$, and that it is nondecreasing in each $n_{i,a}$, so we may also deduce that $e_a(\mathbf{x}_{\mathcal{B}}) \leq e_a(\mathbf{x}_{\mathcal{A}})$ for all a . Hence we consider a parameterized path $p(\tau)$ in $\mathbb{R}^{(m+1) \cdot \ell}$ from $p(0) := (\mathbf{n}(\mathbf{x}_{\mathcal{B}}), \mathbf{e}(\mathbf{x}_{\mathcal{B}}))$ to $p(1) := (\mathbf{n}(\mathbf{x}_{\mathcal{A}}), \mathbf{e}(\mathbf{x}_{\mathcal{A}}))$. Then by integrating along the path we obtain

$$(8.3) \quad \Delta(t \mid \mathbf{x}_{\mathcal{A}}) - \Delta(t \mid \mathbf{x}_{\mathcal{B}}) = \int_{\tau=0}^1 \left(\frac{d(\theta \circ p)}{d\tau} \right) d\tau.$$

We require that at each point in p it holds that $e_a = \frac{1}{2} \sum_{i \neq j} \sum_{b \neq a} n_{i,b} \cdot n_{j,b}$ for all a , and also ensure that p is nondecreasing in each coordinate. There exists a path meeting these requirements, since $(\mathbf{n}(\mathbf{x}_{\mathcal{B}}), \mathbf{e}(\mathbf{x}_{\mathcal{B}})) \leq (\mathbf{n}(\mathbf{x}_{\mathcal{A}}), \mathbf{e}(\mathbf{x}_{\mathcal{A}}))$ and

each e_a is nondecreasing in each $n_{i,b}$ variable. This implies $\partial n_{i,a}/\partial \tau \geq 0$ and $\partial e_a/\partial \tau \geq 0$ for all i and a . Hence we can prove the integral is nonnegative by applying the chain rule for the derivative to obtain

$$\frac{d(\theta \circ p)}{d\tau} = \sum_{i,a} \frac{\partial \theta}{\partial n_{i,a}} \frac{\partial n_{i,a}}{\partial \tau} + \sum_a \frac{\partial \theta}{\partial e_a} \frac{\partial e_a}{\partial \tau}$$

and then proving that $\partial \theta/\partial n_{i,a} \geq 0$ and $\partial \theta/\partial e_a \geq 0$ for all i and a . Next, observe that $\partial \theta/\partial e_a = (1 - n_a/n_V) \geq 0$. So fix a class index k and an outcome c and consider $\partial \theta/\partial n_{k,c}$. Elementary calculus tells us that

$$(8.4) \quad \frac{\partial \theta}{\partial n_{k,c}} = \sum_{j \neq k, b \neq c} n_{j,b} + \sum_b \frac{e_b n_b}{n_V^2} - \frac{e_c}{n_V}$$

This quantity is nonnegative iff

$$(8.5) \quad n_V e_c \leq n_V^2 \cdot \sum_{j \neq k, b \neq c} n_{j,b} + \sum_b e_b n_b$$

Now substitute $\frac{1}{2} \sum_{i \neq j} \sum_{b \neq c} n_{i,b} \cdot n_{j,b}$ for e_c to obtain

$$(8.6) \quad n_V e_c = \frac{n_V}{2} \sum_{i \neq j} \sum_{b \neq c} n_{i,b} \cdot n_{j,b} \leq n_V \cdot \left(\sum_{j \neq k, b \neq c} n_{j,b} \right) \left(\sum_{i,a} n_{i,a} \right) \leq n_V^2 \cdot \left(\sum_{j \neq k, b \neq c} n_{j,b} \right)$$

Since $\sum_b e_b n_b \geq 0$, we obtain Eq. (8.5) from Eq. (8.6) by inspection, and hence $\partial \theta/\partial n_{k,c} \geq 0$ for all k and c . This completes the proof of the adaptive submodularity of f under a uniform prior.

We now show how to reduce the general prior case to the uniform prior case. Fix any prior P with rational probabilities, i.e. $P(h) \in \mathbb{Q}$ for all h . Then there exists $d \in \mathbb{N}$ and function $k : \mathcal{H} \rightarrow \mathbb{N}$ such that $P(h) = k(h)/d$. Create a new instance containing d hypotheses, where for each $h \in \mathcal{H}$ there are $k(h)$ copies of h , denoted by $h^1, \dots, h^{k(h)}$. Each copy of h induces the same conditional distribution of test outcomes $P(X_1, \dots, X_N \mid h)$. All copies of h belong to the same class, and copies of h and h' belong to the same class iff h

and h' do. Finally, assign a uniform prior to this new instance. Then the adaptive submodularity of f on this new instance implies the adaptive submodularity on the original instance, if the weight of edge $\{h, h'\}$ in the original instance is proportional to the number of edges between the copies of h and the copies of h' in the new instance. That is, it suffices to set $w(\{h, h'\}) \propto k(h) \cdot k(h')$, and our choice of weight function, $w(\{h, h'\}) := P(h) \cdot P(h')$, satisfies this condition. *Q.E.D.*

A Bad Example for Information Gain

A popular heuristic for the Optimal Decision Tree problem are to adaptively greedily select the test that maximizes the *information gain* in the distribution over hypotheses, conditioned on all previous test outcomes. The same heuristic can be applied to the Equivalence Class Determination problem, in which we compute the information gain with respect to the entropy of the distribution over *classes* rather than hypotheses. Let π_{IG} denote the resulting policy for Equivalence Class Determination.

Another common heuristic for Optimal Decision Tree is to adaptively greedily select the test maximizing the *Bayesian decision-theoretic value of information* (VoI) criterion. Recall the value of information of a test t is the expected reduction in the expected risk of the minimum risk decision, where the risk is the expected loss. Formally, consider the Bayesian decision-theoretic setup described in §3.3. The VoI criterion myopically selects test to maximize

$$\Delta_{\text{VoI}}(t | \mathbf{x}_{\mathcal{A}}) := \min_d \mathbb{E}_H[\ell(d, H) | \mathbf{x}_{\mathcal{A}}] - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_{\mathcal{A}}} \left[\min_d \mathbb{E}_H[\ell(d, H) | \mathbf{x}_{\mathcal{A}}, x_t] \right].$$

This heuristic can be also be applied to the Equivalence Class Determination problem, by taking the decision set \mathcal{D} to be the set of equivalence classes, and the loss function to be the 0–1 classification loss function, i.e., $\ell(d, H) = \mathbf{1}[H \notin d]$. Let π_{VoI} denote the resulting policy.

In this section we present a family of Equivalence Class Determination instances for which both π_{IG} and π_{VoI} perform significantly worse than the optimal policy.

THEOREM 7 *There exists a family of Equivalence Class Determination instances with uniform priors such that $c(\pi_{IG}) = \Omega(n/\log(n)) c(\pi^*)$ and $c(\pi_{VoI}) = \Omega(n/\log(n)) c(\pi^*)$, where n is the number of hypotheses and π^* is an optimal policy.*

In fact, we will prove a lower bound for each policy within a large family of adaptive greedy policies which contains π_{IG} and π_{VoI} , which we call *posterior-based*. Informally, this family consists of all greedily policies that use only information about the posterior equivalence class distribution to select the next test. More precisely, these policies define a potential function Φ which maps distributions of distributions over equivalence classes to real numbers, and at each time step select the test t which maximizes Φ of the posterior distribution (over test outcomes x_t) of the posterior distribution over equivalence classes generated by adding x_t to the previously seen test outcomes. In the event of a tie, we select any test maximizing this quantity at random. The information gain policy is posterior-based; Φ is simply -1 times the expected entropy of the posterior equivalence-class distribution. Likewise, the value of information policy is also posterior-based; Φ is simply -1 times the expected loss of the best action for the posterior equivalence-class distribution. Hence to prove Theorem 7 it suffices to prove the following more general theorem.

THEOREM 8 *There exists a family of Equivalence Class Determination instances with uniform priors such that $c(\pi) = \Omega(n/\log(n)) c(\pi^*)$ for any posterior-based policy π , where n is the number of hypotheses and π^* is an optimal policy.*

PROOF: Fix integer parameter $q \geq 1$. There are $m = 2^q$ classes \mathcal{H}_a for each $1 \leq a \leq 2^q$. Each \mathcal{H}_a consists of two hypotheses, $h_{a,0}$ and $h_{a,1}$. We call a the in-

dex of \mathcal{H}_a . The prior is uniform over the hypotheses $\mathcal{H} = \{h_{a,v} : 1 \leq a \leq m, 0 \leq v \leq 1\}$. There are four types of tests, all with binary outcomes and all of unit cost. There is only one test of the first type, t_0 , which tells us the value of v in the realized hypothesis $h_{a,v}^*$. Hence for all a , $H = h_{a,v} \Rightarrow X_{t_0} = v$. Tests of the second type are designed to help us quickly discover the index of the realized class via binary search if we have already run t_0 , but to offer no information gain whatsoever if t_0 has not yet been run. There is one such test t_k for all t with $1 \leq k \leq q$. For $z \in \mathbb{N}$, let $\phi_k(z)$ denote the k^{th} least-significant bit of the binary encoding of z , so that $z = \sum_{k=1}^{\infty} 2^{k-1} \phi_k(z)$. Then for each $h_{a,v}$ we have $H = h_{a,v} \Rightarrow X_{t_k} = 1[\phi_k(a) = v]$. Tests of the third type are designed to allow us to do a (comparatively slow) sequential search on the index of the realized class. Specifically, we have tests t_k^{seq} for all $1 \leq k \leq m$, such that $H = h_{a,v} \Rightarrow X_{t_k^{\text{seq}}} = 1[a = k]$. Finally, tests of the fourth type, $\{t_k^{\text{dumb}} : k \in \mathbb{N}\}$, are dummy tests that reveal no information at all. Formally, $X_{t_k^{\text{dumb}}}$ always equals zero.

Given this input, suppose $H = h_{a,v}$. One solution is to run t_0 to find v , then run tests t_1, \dots, t_q to determine $\phi_k(a)$ for all $1 \leq k \leq q$ and hence to determine a . This reveals the value of H , and hence the class H belongs to. Since the tests have unit cost, this policy π' has cost $c(\pi') = q + 1$.

Next, fix a posterior-based policy π and consider what it will do. Call a class *possible* if not all of its hypotheses have been ruled out by tests performed so far. Note that all possible classes contain the same number of hypotheses, because they initially have two, and each test t_k that can reduce the size of a possible class to one, will reduce the size of every possible class to one. This, and the fact that the prior is uniform, implies that the posterior equivalence-class distribution is uniform over the remaining possible classes. If no tests in $\{t_k : 0 \leq k \leq q\}$ have been run, as is initially the case, any single test in this set will not change the posterior equivalence-class distribution. Hence, as measured with respect to Φ , such tests are precisely as good as the dummy tests. If these tests are each better

than any test in $\{t_k^{\text{seq}} : 1 \leq k \leq m\}$, then π selects among $\{t_k : 0 \leq k \leq q\} \cup \{t_k^{\text{dumb}} : k \in \mathbb{N}\}$ at random. Since there are infinitely many dummy tests, with probability one a dummy test is selected. Since the posterior remains the same, π will repeatedly select a test at random from this set, resulting in an infinite loop as dummy tests are selected repeatedly *ad infinitum*. Otherwise, some test t_k^{seq} is preferable to the other tests, measured with respect to Φ . In the likely event that t is not the index of H , we are left with a residual problem in which tests in $\{t_k : 0 \leq k \leq q\}$ still have no effect on the posterior, there is one less class, and the prior is again uniform. Hence our previous argument still applies, and π will either enter an infinite loop or will repeatedly select tests in $\{t_k^{\text{seq}} : 1 \leq k \leq m\}$ until a test has an outcome of 1. Thus in expectation π costs at least $c(\pi) \geq \frac{1}{m} \sum_{z=1}^m z = (m+1)/2$. Since $m = 2^q$, $n = 2m$, and $c(\pi^*) \leq c(\pi') = q+1 = \log_2(n)$ we infer

$$c(\pi) \geq \frac{m}{2} = \left(\frac{n}{4 \log_2(n)} \right) c(\pi^*)$$

which completes the proof.

Q.E.D.